Genetic Basis of Response to 50 Generations of Selection on Body Weight in Inbred Mice

Peter D. Keightley

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland

Manuscript received September 18, 1997
Accepted for publication December 11, 1997

ABSTRACT

A long-established inbred strain of mice was divergently selected for body weight for 50 generations. Selection of new mutations affecting the trait eventually led to a divergence of approximately three phenotypic standard deviations between the high and low lines. Heritability for body weight increased at a rate between 0.23% and 0.57% per generation from new mutations, depending on the genetic model assumed. About two-thirds of the selection response was in the upward direction. The response was episodic, suggesting a substantial contribution from the selection of mutations with large effects on the trait. A maximum likelihood procedure was used to estimate the number of factors contributing to the response using data from line crosses, with models of n equivalent gene effects (*i.e.*, to estimate the Wright-Castle index), or n genes with variable effects. The results of the analysis of data from a cross between the selected high line and an unselected control line indicated that two major factors were involved, with the suggestion of an additional minor factor.

THE rate of increase of genetic variance of quantitative traits from the accumulation of new mutations has been known for some time to be on the order of 0.1% of the environmental variance per generation (Clayton and Robertson 1955; Falconer and Mackay 1996). Theory to predict the contribution of new mutations to response to artificial selection has been developed by Hill (1982a,b). The contribution of mutational variation to the response is very sensitive, particularly in the short term, to the nature of the mutational variation. At one extreme, the infinitesimal model of many unlinked, mutant alleles—each with small additive effects—predicts a nonlinear, accelerating pattern of response, which eventually reaches an asymptotic rate of $R = 2NiV_m/\sigma_p$, where N is the effective population size, i is the selection intensity, V_m is the increment in variance from one generation of mutation, and σ_p is the phenotypic standard deviation. At the opposite extreme, if mutational variation is contributed by a small number of mutations with large additive effects, the predicted asymptotic rate of response is the same as under the infinitesimal model, but the asymptotic rate is expected to be reached more quickly. For both models, the response reaches a rate proportional to the effective population size, and this is an argument for maintaining commercial selection lines at as large as possible population sizes. Response from mutations with large effects is expected to be highly variable, as it depends on their chance appearance and fixation, but the response is linearly related to population size because fixation probability is essentially independent of N, but the number of mutation events is proportional to N. Recessive mutant alleles are expected to make small contributions to initial selection response, as their initial rates of frequency change are slow.

Two experiments in Drosophila to measure the contribution of new mutations to selection response for bristle number using inbred or isogenic base populations have been carried out on a large enough scale to allow the evaluation of some of the theoretical predictions just described. Lopez and Lopez-Fanjul (1993a) divergently selected for 47 generations on abdominal bristle number in replicated lines with two different population sizes. As predicted from theory, the total response was approximately proportional to the population size selected. The response averaged over replicates for both population sizes was nonlinear, but not strongly so. The highly variable and episodic appearance of the response indicated that mutations with large effects were becoming fixed or brought to high frequency by selection. Subsequent genetic analysis of the lines also suggested an important role for mutations with large effects, particularly from deleterious recessives (Lopez and Lopez-Fanjul 1993b). Mackay *et al.* (1994) divergently selected for abdominal and sternopleural bristle number for 125 generations and observed highly variable selection responses among replicates, which were episodic in appearance, suggesting the selection of mutations with large effects. Responses were also highly asymmetrical, and there were correlated reductions in fitness (Nuzhdin *et al.* 1995). A line-cross analysis to estimate

Address for correspondence: Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland. E-mail: p.keightley@edinburgh.ac.uk

the effective number of loci involved suggested that several loci had contributed to the response in each line (Fry *et al.* 1995).

Most of our knowledge of the potential for new mutations to contribute to artificial selection responses is restricted to Drosophila bristle number. In mice, there are a number of estimates of the rate of accumulation of variance for various morphological traits, based on rates of divergence between inbred sublines, which suggests that heritability increases at least an order of magnitude faster than is typical for Drosophila morphological traits (reviewed by Houle *et al.* 1996). However, the validity of the estimates for the mouse has been questioned because sublines were measured at different times in different laboratories (Caballero *et al.* 1995). There is essentially no information on the genetic basis of mutational variability for quantitative traits in mammals. This article reports results of experiments to investigate the nature of the selection response from new mutation in lines of mice that have been divergently selected for body weight for 50 generations from an inbred base population. The effective number of factors contributing to the response is estimated using a maximum likelihood (ML) approach to infer the Wright-Castle index, and the approach is extended to estimate the minimum number of factors required to explain the data with a model of variable gene action. Previous estimates for rates of accumulation of mutational variability for body weight in the mouse are updated.

MATERIALS AND METHODS

Mouse selection line and mutational variance estimation: The origin and maintenance of the lines and the methods used to estimate mutational variance have been described previously (Keightley and Hill 1992; Caballero *et al.* 1995). Briefly, high and low selection lines for 6-wk body weight were established in 1986 from a subline of the inbred strain C3H/He, which had previously been maintained by brother-sister mating for more than 160 generations (Festing 1989). Selection was on a within-family basis. Up to generation 21, a circular mating scheme was used (Kimura and Crow 1963), but it was discontinued in favor of a scheme in which half the matings were between full sibs and the remainder were at random, to increase the probability of selection of recessive mutations. At generation 37, the C3H/He inbred was obtained from the original supplier (Bantin and Kingman Ltd., Hull, England), which had continued to maintain the line by brother-sister mating. This line was used as an unselected control, with six matings per generation and the same mating scheme as for the selection lines. Assuming that the inbred had remained genetically constant, contemporary measurement of its body weight allows the direction of any selection response in the lines to be determined. At generation 36, the mouse lines were moved by embryo transfer to a new mouse house in which pathogen levels were lower and the environmental conditions more constant. The mice were maintained at 21° on the same standard diet throughout the experiment (rat and mouse no. 3 breeding diet to 3 wk of age, no. 1 maintenance diet thereafter; SDS Ltd., Essex, England). The data from the selection experiment comprise records on 6993 mice in 1434 litters (634 low-line and 702 high-line litters, and an additional 98 control-line litters). The selection experiment has been terminated at generation 50, and the lines are being re-inbred.

The increment of genetic variation per generation, V_m, was estimated in three different ways:

1. By an animal-model restrictd maximum likelihood (REML) analysis that assumes the infinitesimal model of many unlinked genes with small additive effects and uses all the data in the pedigree (Meyer 1989). The genetic variation in the inbred line was assumed to be $4V_m$, as expected for a line at mutation-drift balance maintained by brother-sister mating (Lynch and Hill 1986). Previously (*e.g.*, Keightley and Hill 1992), the base population variance has been erroneously assumed to be $5V_m$, which leads to marginally different estimates. A random litter effect and fixed effects of sex, generation number, and litter size were included in the model.
2. By fitting the expected response under the infinitesimal model to the observed high-low, high-control, or control-low divergences by least squares using the equation of Hill (1982b).
3. By fitting the expected response under a model of additive genes with large effects fixed rapidly by selection to the observed divergences by least squares, again using the equation of Hill (1982b).

Both infinitesimal and large-gene-effects models assume that mutations affect fitness only through their effects on the artificially selected trait. In fitting the expected divergences to the observed high-low divergence, a nonzero intercept was fitted to account for a response induced by a maternal effect. With data from the high-control or control-low divergences, a zero intercept had to be assumed, as there were no data in the initial generations to reliably estimate the intercept. The mutational variance is scaled relative to the environmental variance between and within letters and is expressed as the mutational heritability, $h_m^2 = V_m/V_e$.

Relaxed selection lines and crosses between them: At generation 43, sublines were split from the high and low selection lines and maintained without selection for two generations with 10-pair matings. The control line was also maintained at 10-pair matings in generations 44 and 45. To investigate the genetic basis of the upward selection response that had occurred by this time, the progeny of the relaxed high line from generations 44 and 45 were crossed to the control line, and 6-wk body weights of large populations of F_1 and F_2 mice were recorded.

ML estimation of Wright's effective number of loci: The number of loci, n, contributing to the difference, R, between divergently selected lines can be estimated using Wright's formula, which relates R to the genetic variance when the lines are crossed (Falconer and Mackay 1996, chapter 12) and assumes (1) that there are n unlinked genes, each with equal additive effects, a (half the difference between homozygotes) and (2) that alleles are fixed in opposite directions in the divergent selection lines. Here, a dominance term d (the difference between the heterozygote and the mean of the homozygotes), assumed to be equal for all loci, is also estimated. Estimation of additive and dominance terms requires data from the parental lines, the F_1 and the F_2, so it seemed most appropriate to analyze all the data together by ML, which allows comparison of the fit of the data to the models with different n. As well as the genetic effects, a normally distributed effect common to full-sib litter mates, a normally distributed individual environmental effect, fixed effects of sex, parity of birth, and generation (each with two levels), and a linear covariate effect for litter size were simultaneously fitted. The

following method and nomenclature are based on Haley *et al.* (1993), who applied ML to test for a single major gene in a population subdivided into families with background effects common to families. The likelihood of data from parental, F_1 and F_2 populations can be written

$$L = \prod_{i=1}^{N_L} \int \frac{1}{(2\pi\sigma_b^2)^{1/2}} exp\left[-\frac{b_i^2}{2\sigma_b^2}\right] \prod_{j=1}^{m_i} \sum_{g=1}^{states} freq_c(g) \frac{1}{(2\pi\sigma_w^2)^{1/2}} exp\left[-\frac{(y_{ij} - \mu - a_g - d_g - b_i - \mathbf{x}_{ij}'\mathbf{f})^2}{2\sigma_w^2}\right] db_i, \quad (1)$$

where N_L is the number of litters, m_i is the number of individuals within litter i, b_i is the random effect of the ith litter, assumed to be normally distributed with mean zero and variance σ_b^2, y_{ij} is the observed trait value of individual j from litter i, $freq_c(g)$ is the frequency of the multilocus genotype g from line c from the $states = 3^n$ possible genotypes for the n genes, μ is the population mean, a_g and d_g are additive and dominance effects, respectively, for genotype g (the sum of the additive or dominance effects of each locus), $\mathbf{x}_{ij}$ is the design matrix for fixed effects and the covariate for individual ij, $\mathbf{f}$ is the vector of fixed and covariate effects and σ_w^2 is the residual variance. The parental lines are assumed to be fixed for opposite alleles acting in the same direction at all the loci at which they differ, and the F_1 is assumed to be heterozygous at all loci, so for these subpopulations there is only one nonzero $freq_c(g)$, which takes the value one. For the F_2, the values of $freq_c(g)$ were obtained from a binomial expansion. For a given n, likelihood was computed by evaluating Equation 1 using Simpson's rule to numerically integrate for the litter effects. Maximization as a function of the parameter values was performed using the simplex algorithm (Nelder and Mead 1965), and convergence to the ML was checked by restarting the procedure after convergence to a maximum until no further significant increase in likelihood was found, a strategy to avoid spurious convergence (Press *et al.* 1992). For each n, several runs were performed with widely different starting values to explore the possibility of local maxima of the likelihood, but only putative global maxima were found.

Variable gene effects: The classic model to estimate the number of loci explaining fixed differences between selection lines assumes equal gene effects for all the n loci affecting the trait. By ML, is is also possible to compare the fit of the model for cases of small numbers of loci, each with different additive and dominance effects, by evaluating Equation 1. Each additional variable locus introduces two extra parameters, compared with the model with equal effects. The models therefore have large numbers of parameters, so the multidimensional likelihood surfaces were explored extensively to check for local maxima, but none were found for cases of one, two, or three variable loci. With four variable loci, difficulty was encountered in locating the global maximum, as there appeared to be a local maximum. The likelihood space was explored from 12 widely differing combinations of starting parameter values and was found in all cases to converge either to the one local or to the putative global maximum. Likelihood maximization was also attempted with the Metropolis algorithm with simulated annealing (Press *et al.* 1992, chapter 10), but convergence to either the local or putative global maximum also occurred, and the outcome depended on the choice of starting values.

RESULTS

Response to selection and estimates of mutational variance: Mean 6-wk body weights for the selection and control lines are plotted against generation number in Figure 1. The selection response was episodic, suggesting the appearance and fixation of a small number of major mutations. The bursts of response are more obvious in the plot of the high-low divergence (Figure 2), which also shows that there was little response in the last 12 generations. About two-thirds of the response was upward and possibly associated with a jump between generations 35–40 (Figure 2), although the lack of control line data before generation 37 makes it difficult to tell whether the jump occurred in the low or high line. The greater upward than downward response is surprising, because in mice, very many more mutations are known that reduce growth than increase it (Lyon and Searle 1989).

The episodic nature of the response implies that the

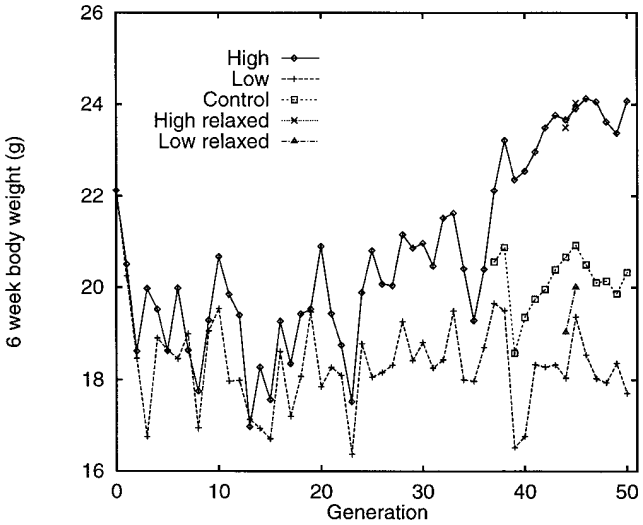

Figure 1.—Mean 6-wk body weight (g) corrected for litter size, averaged over sexes, for high and low selection lines and the control lines from generation 37, and high and low relaxed selection lines in generations 44 and 45.

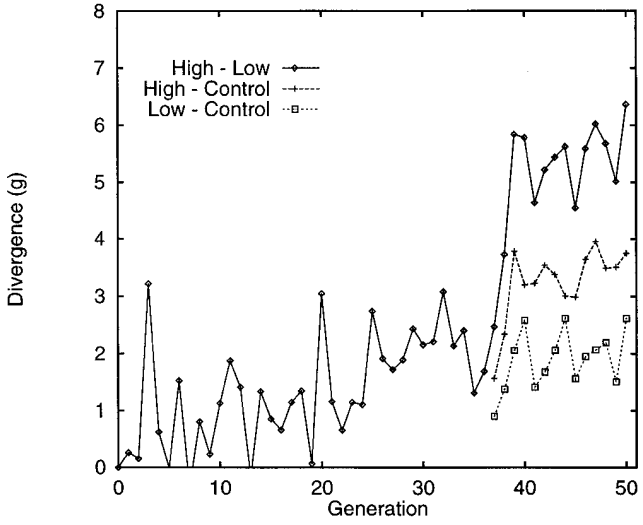

Figure 2.—Mean divergences for 6-wk body weight (g) plotted against generation number.

TABLE 1

Estimates of mutational heritabilities for 6-wk body weight

Model	Data	h_m^2	Support limits
Infinitesimal (REML)	Complete pedigree	0.0053	0.0038–0.0072
Infinitesimal	High-low divergence	0.0057	
Large gene effects	High-low divergence	0.0023	
Infinitesimal	High-control divergence	0.0070	
Large gene effects	High-control divergence	0.0025	
Infinitesimal	Control-low divergence	0.0056	
Large gene effects	Control-low divergence	0.0023	

infinitesimal model of many additive, unlinked mutations contributing to the response is inappropriate. An alternative model, the "large-gene-effects model," assumes unlinked, additive mutations of large effect, which become fixed in a short time scale relative to the duration of the experiment, but this also has the drawback of assuming additive mutation effects. Nonetheless, these two models are standard benchmarks for quantifying the mutational input for quantitative traits. Estimated rates of increase of heritability from mutation per generation, (h_m^2), based only on information from the divergence between the high and low lines or between the selection lines and control line, assuming either the infinitesimal or large-gene-effects model, are compared to an estimate from the animal-model REML analysis in Table 1. The animal-model REML analysis assumes the infinitesimal model, with mutational variation incorporated in the numerator relationship matrix (Wray 1990), and uses information from covariances between relatives as well as information from the response. The mutational heritability estimates from the infinitesimal model high-low divergence and animal-model REML are both close to 0.5% (Table 1). With the animal model, the information to estimate heritability comes increasingly from the response as generation number increases (Juga and Thompson 1989), and this presumably explains why the estimates are close. The mutational heritability estimate from the high-low divergence under the large-gene-effects model is 0.23%. With large mutant effects, a lower estimate is expected because predicted initial rates of response for a given input of mutational variation are higher, but predicted *asymptotic* rates are similar for both models (Hill 1982a,b). Mutational heritability estimates from the high-control and control-low divergences are higher than the estimate from the high-low divergence (Table 1) because the fitted response curves are constrained to pass through the origin in the former case. The h_m^2 estimates from high-control and control-low are not proportional to the observed divergences (Figure 2), principally because the realized selection intensity was lower in the low line than in the high line (0.33 *vs.* 0.44).

Relaxed selection: During the first 20 generations of the experiment, the high and low lines differed, on average, by 1.0 g (Figure 2), a likely consequence of a maternal effect induced by selection. A maternal effect is expected because high (low) selected mothers will tend to have larger (smaller) offspring than average. This effect is expected to disappear rapidly if unselected parents are used (Kirkpatrick and Lande 1989). To further investigate the nature of the selection responses, lines were maintained without selection for two generations. A reversal of the selection response occurred in the low line (Figure 1): there is a difference of 0.8 g in mean body weight between the relaxed low line and the selected low line, but mean phenotypic values for mice from the relaxed high line were very close to values from the selection lines. The magnitude of the change in mean body weight after relaxation in the low line is therefore similar to the maternal effect observed in the initial generations of the experiment. Alternatively, the reversal in the low line after relaxation could have been caused by the segregation of deleterious mutations maintained at intermediate frequencies by selection on the trait.

Means and variances in a cross between high and control lines and calculation of effective number of loci: Means and variance components from REML analysis (Genstat 5 Committee 1993) of data from generations 44–45 of the relaxed selection lines, the control line, and their F_1 and F_2 are compared in Table 2. The difference between the relaxed high and control lines was 3.46 g (more than two phenotypic SD), while the difference between the control and relaxed low line was only 0.3 g, but the latter difference is somewhat increased if the data are corrected for litter size (Figure 2). The F_1 is closer in body weight to the high line, suggesting dominance of high-line mutant alleles. By generation 44, the response appeared to have reached a plateau, so it is reasonable to assume that mutant alleles are fixed in the high line and that the difference in within-litter variance between the F_2 and the F_1 can be equated to the genetic variance, V_g. The effective number of loci, n, assuming equal additive and dominance effects for each locus, is related to the high-control difference, R, and the deviation from the F_1 from the mean of its parents, D, according to $n = (R^2/8 + D^2/4)/V_g$. Substi-

TABLE 2

Mean 6-wk body weights (g) and variance components (g^2) from REML analysis of data from generations 44 and 45 from the relaxed low and high lines, the control lines, and the F_1 and F_2

Line or cross	No. of mice	$\bar{X}$	σ_b^2	σ_w^2
Low relaxed	99	19.9 (0.33)	0.43 (0.27)	1.57 (0.25)
High relaxed	134	23.6 (0.32)	0.48 (0.25)	1.49 (0.20)
Control	172	20.2 (0.38)	1.04 (0.44)	1.41 (0.16)
F_1 high × control	315	22.5 (0.21)	0.50 (0.16)	1.26 (0.11)
F_2 high × control	462	22.2 (0.17)	0.68 (0.20)	2.24 (0.15)

Values in parentheses indicate SE.

tution of the observed R (3.46 g), D (0.61 g) and V_g (0.98 g^2) values (Table 2) gives an estimate for n of 1.6. This estimate is consistent with the pattern of selection response, which suggests that one or two major loci were involved.

Estimation of number of loci differentiating the lines: The ML procedure detailed in materials and methods was applied to the complete, untransformed data set from the high and control lines and to their F_1 and F_2 from generations 44 and 45. Natural log likelihood of the data as a function of the number of loci, under a model of equal additive and dominance effects, is shown in Figure 3. A two-locus model gives the best fit to the data, with a considerably higher likelihood than the one-locus model (the likelihood ratio is $e^{6.3} = 545$). A significance test is not possible, however, as the constraint of equal additive and dominance effects implies that the definition of the parameters changes as extra loci are added to the model. To allow such tests, models in which loci have variable additive and dominance effects were also investigated (Table 3). Likelihood must increase as extra loci are added to the model. Likelihood for two variable loci turned out to be the same as for two equivalent loci: two equal loci maximize the variance for a given difference between the line means, and this is presumably the major factor determining the fit. The addition of one locus with an additive and a dominance parameter implies that the change in twice log likelihood follows a chi-square distribution with two degrees of freedom, so the change in log likelihood of 6.3 between the one- and two-locus models is significant ($P < 0.01$). Somewhat surprisingly, the addition of a third variable locus also resulted in a significant increase in log likelihood ($P < 0.01$). The best fitting three-locus model was two major dominant loci and one minor, underdominant locus (Table 3), a result which is difficult to explain intuitively. With four variable loci, there were two maxima in the likelihood surface, the first with three dominant loci and an underdominant locus, and the second with two dominant loci and two underdominants, and maximization to one or the other of these occurred, depending on the initial parameter values. The four-locus model with two underdominant loci gave a higher likelihood, however, but not significantly higher than the three-locus model (Table 3).

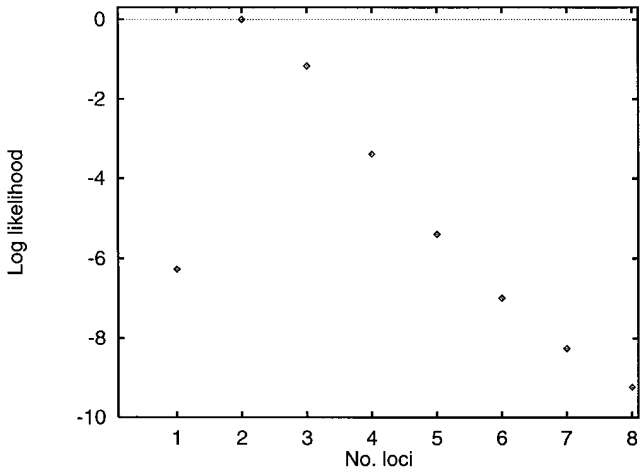

Figure 3.—Natural log likelihood of data from control and relaxed high lines from generations 44 and 45, the F_1 between them from generations 44 and 45, and their F_2 from generation 45, as a function of number of fixed loci.

Realized selection differential: The pattern of the selection response, particularly the divergence between the high and low lines (Figure 2), suggests that the response had reached a plateau after generation 40. Natural selection opposing artificial selection because of selection of alleles with deleterious pleiotropic effects on fitness is one common explanation for selection limits (Falconer and Mackay 1996, chapter 12). However, in the present experiment there was almost no detectable change in either litter size or viability (Caballero and Keightley 1998). To further explore the fitness effects of the selection on body weight, the realized selection differential was calculated and is shown for the high and low lines separately for males and females in Figure 4. Although the selection differential varied considerably from generation to generation, there is no indication of a loss of selection intensity in the high line (it actually increased slightly, by 0.03 g over the 50 generations). In the low line, there is a slight indication of a loss of selection differential, as it dropped by about 0.002 g per generation, averaged over sexes. Regression coefficients

TABLE 3

ML estimates of additive and dominance effects of genes with a model of unequal effects

No. of loci	Estimated gene effects (g)								Relative Log_eL
	$\hat{a}_1$	$\hat{d}_1$	$\hat{a}_2$	$\hat{d}_2$	$\hat{a}_3$	$\hat{d}_3$	$\hat{a}_4$	$\hat{d}_4$	
1	2.6	1.8							0.0
2	1.6	1.1	1.6	1.1					6.3
3	1.5	1.3	1.5	1.3	0.0	−0.7			10.1
4	1.4	1.4	1.4	1.4	0.1	−0.5	0.1	−0.5	10.9

calculated within each sex for selection differential on generation number were nonsignificantly different from zero ($P > 0.2$). The more likely explanation for the plateau is a lack of genetic variation in the lines.

Age-specific effects on body weight: To test for age-specific differences in body weight between the lines, additional measurements were taken at 3 and 10 wk of age at generation 47 (Figure 5). Absolute differences between the high and control lines increase with age, but the relative differences are highest at 3 wk (35%, dropping to 20% at 6 and 10 wk).

Checks on contamination of selection lines: With the exception of a phaeomelanin-deficient mutant *rimy* (Keightley and Hawkins 1991), which reduced body weight and so segregated briefly in the high line, the coat-color phenotype of every mouse recorded was wild type. At generation 36, all mice that contributed offspring to the next generation were typed with a nonecotropic, retrovirus-specific probe by Southern blotting (Keightley and Bulfield 1993). These mice showed identical retrovirus fingerprints, while control mice (C57BL/6 and DBA/2) showed many differences (data not shown). At generation 43, a sample of 23 mice from the high, low, and control lines was typed at 10 unlinked *Mit* microsatellite loci (Dietrich *et al.* 1992). PCR products, which were separated on 20-cm polyacrylamide gels, appeared to be invariant in size (data not shown), again suggesting that the marker loci are monomorphic in these individuals.

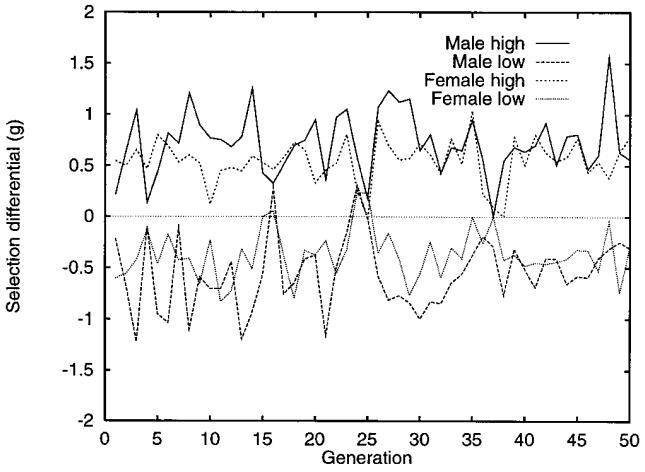

Figure 4.—Mean selection differential (g) for males and females in the high and low lines, plotted separately as a function of generation number.

DISCUSSION

Genetic variation for body weight from new mutations: The mutational heritability estimate reported here of 0.53% from the animal-model REML analysis is about half the value reported from this experiment at generation 24 (Keightley and Hill 1992)—which is surprising, given that almost all of the response occurred after generation 24. With increasing generation number, information to estimate h_m^2 comes increasingly from the response to selection, rather than from covariances between close relatives (Juga and Thompson 1989). The high early h_m^2 estimate in the absence of a response could be explained by a buildup of deleterious mutant alleles with pleiotropic effects on body weight, which could not be subsequently fixed (Caballero *et al.* 1995). However, the lack of a downward drift of fertility or viability and the essential absence of change in the realized selection differential (Figure 4) tend to

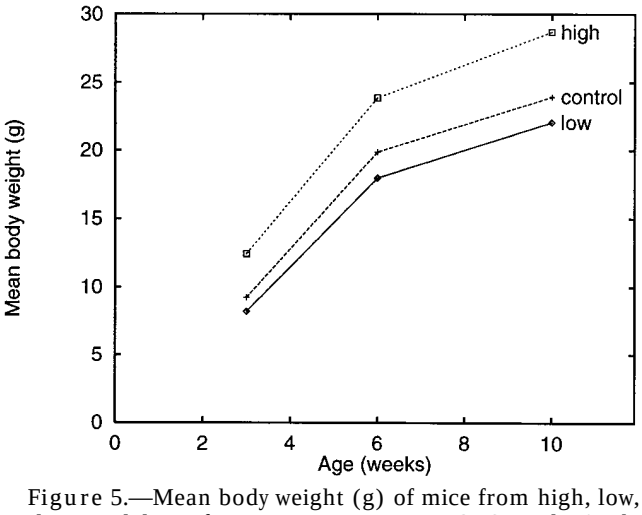

Figure 5.—Mean body weight (g) of mice from high, low, and control lines from generation 47 at 3, 6, and 10 wk. Standard errors of differences between the body weights for the three lines at the same age are about 0.6 g.

suggest that this is not the correct explanation. More likely, perhaps, is the presence of uncontrolled environmental factors (*e.g.*, disease) leading to nongenetic resemblances between relatives. Confounding factors of this nature would become less important as increasing information comes from the selection response. The close agreement between the h_m^2 estimates under the infinitesimal model from the animal-model REML analysis and the analysis using the response only is reassuring in this respect. At generation 24, the REML- and divergence-based h_m^2 estimates disagreed (Keightley and Hill 1992). It can be argued that because the response appeared to reach a plateau and was episodic, the lower h_m^2 estimate from the large-gene-effects model is more meaningful. This figure is about twice the figure from a different experiment to estimate h_m^2 for body weight in mice involving selection in a cross of two long-separated inbreds (Caballero *et al.* 1995). Under the infinitesimal model (the more usual model), the estimated mutational heritability is somewhat higher than typically found for morphological traits in animals (Houle *et al.* 1996; most information is for bristle number in Drosophila: an estimate for h_m^2 for a comparable trait of Drosophila, wing length, is 0.2%), but is very much lower than estimates for a variety of skeletal traits in mice based on divergences of inbred sublines. It is a strong possibility that the mutational components of variance for skeletal traits have been overestimated, because in many cases the phenotypes were measured at different times in lines maintained in different environments (Festing 1973; Caballero *et al.* 1995).

The extent of the response seen in the present experiment lends support to Hill's (1982a,b) suggestion that new mutations can make large contributions to responses in breeding programs. There are several examples of jumps in selection responses in mouse selection experiments for body size involving outbred lines (Roberts 1966; Bradford and Famula 1984; Heath *et al.* 1995), presumably attributable to selection of newly arisen mutations, although rare recombination events are also a possibility. In the present experiment, the population size was small and the selection intensity weak because of the small average family size of the inbred strain. Long-term responses from new mutations are expected to be proportional to the product of selection intensity and effective population size (Hill 1982a,b).

Number of mutations differentiating the lines: It has been emphasized repeatedly in the literature that estimates of the effective number of factors, *n*, tend to be biased downward because the basic method assumes unlinked genes with equivalent effects fixed for favorable alleles in the two lines (see, *e.g.*, Falconer and Mackay 1996, chapter 12). A number of improvements to the basic procedure have been suggested, however. Lande (1981) generalized Wright's procedure to estimate *n* for cases involving genetically heterogeneous parents, F_1, F_2, and backcross generations. Cockerham (1986) suggested that information from individuals from the different generations should be combined and analyzed simultaneously by, for example, least squares, although he did not endorse the use of the procedure. Following simulation studies that highlighted the bias induced by assuming equal gene effects (Zeng *et al.* 1990), Zeng (1992) suggested the inclusion of a composite parameter to account for variability of gene effects in the analysis. The present study has attempted to incorporate these improvements. Likelihood is used to compare the fit of different *n* by analyzing data from all generations simultaneously; fixed and random effects are included in the model; and a dominance term is estimated. The issue of variable gene effects is addressed by comparing the fit of models in which the dominance and additive effects for each locus are allowed to vary. With this model, the analysis is more akin to a segregation analysis, and its use in this context is as an indicator for the presence of major genes.

The results from analysis of the line-cross data by ML point to the response in the high line having been caused by two major mutations, possibly with additional minor mutations. However, the pattern of the selection response seems to show a rapid divergence between the lines at about generation 38, suggesting the fixation of one mutation with a very large effect. Simultaneous fixation of two mutations at about generation 35 seems unlikely. A possible explanation for the discrepancy between the statistical analysis and the response pattern is that segregation analysis methods are known not to be robust to departures from a normal distribution (Go *et al.* 1978; Elston 1979). The ML procedure assumes normally distributed environmental effects, but the distributions of residuals after correction for estimated litter and sex effects showed significant negative skewness ($P < 0.01$; Figure 6). Negative skewness in the F_2 could have a genetic explanation such as segregation of a dominant gene. However, skewness in the control, the high line, and the F_1 (which is attributable to a single outlier) violates the assumption of the model. Unfortunately, power transformations of the raw data (Sokal and Rohlf 1995, chapter 13) did not produce significant improvements in the distribution of residuals.

Nature of mutational variability for body weight: Most information on the nature of spontaneous mutational variation comes from experiments involving selection or random accumulation of mutations in inbred lines of Drosophila, and it is relevant to consider these experiments in relation to the present one. There are large-scale selection experiments for abdominal or sternopleural bristle number in inbred Drosophila, typically resulting in the selection of mutations with large effects on the trait, but very often these mutations are recessive lethals or have detrimental effects on fitness. The mutation effects have been analyzed by chromosome extraction to test for lethals with effect on bristles in the

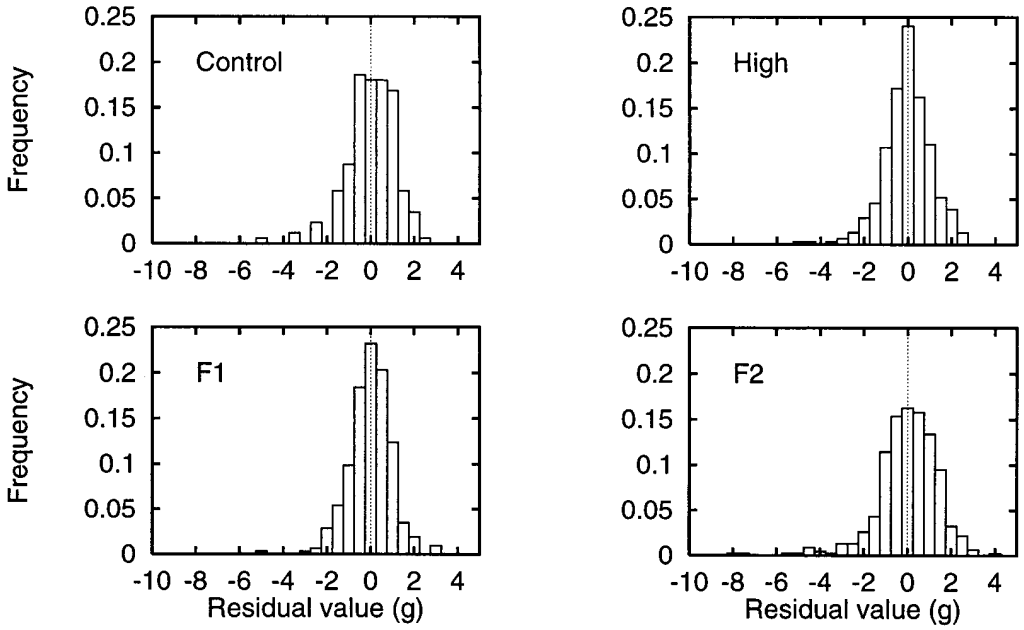

Figure 6.—Frequency distributions of residual values for 6-wk body weight after correction for litter size, sex, the population mean for the control and high lines at generations 44 and 45, and their F_1 and F_2. Values for g_1 are −0.62, −0.72, −0.43, and −1.29, respectively.

heterozygote (Caballero *et al.* 1991; Lopez and Lopez-Fanjul 1993b; Merchante *et al.* 1995; Fry *et al.* 1995), by assays for fitness in the lines themselves (Merchante *et al.* 1995; Nuzhdin *et al.* 1995), or from reversal of response after relaxation of selection, which occurs if there are negative fitness effects of genes with beneficial effects on the trait (Mackay *et al.* 1994). In the present experiment, deleterious fitness effects did not seem to accumulate, as fertility of parents and viability of offspring did not show substantial directional trends, except for a small reduction in fertility of low line mothers whose parents were full sibs and a small reduction of viability in later generations in the low selection line (Caballero and Keightley 1998). Furthermore, relaxation of selection for two generations did not lead to an obvious reversal of the selection response. There are two likely explanations for the absence of negative pleiotropic effects in the present experiment:

1. Artificial selection was weak, corresponding to a within-family selection intensity of only 0.4 standard deviations, so lethal or highly deleterious alleles would be quickly eliminated unless they also had a very large effect on the trait. The weakness of the artificial selection was due to the small family size of the inbred strain.
2. During most of the experiment, half the matings were between full sibs and the remainder between random nonfull sibs. This would have the effect of exposing deleterious recessives to selection and would lead to their rapid elimination. A bristle-number selection experiment in Drosophila with a similar mating scheme gave a substantially lower rate of accumulation of deleterious mutations affecting the trait than a parallel experiment with random mating (Merchante *et al.* 1995).

The general pattern of the selection response was very similar to patterns seen in selection experiments for bristle number in inbred Drosophila, which tend to show periods of apparent stasis punctuated by jumps. In the present experiment, most of the response was for increased body weight and was probably caused by a small number of mutations with large effects, but between generations 38 and 50, little subsequent selection response occurred. There is little evidence that this plateau was caused by a loss of selection intensity because of segregation of deleterious alleles with effects on the trait. A lack of variation in the lines seems to be the more likely explanation for the plateau. Under the infinitesimal model, application of the formula for response from new mutations of Hill (1982b) gives a predicted divergence per generation of about 0.16 g, so the predicted divergence would be about 2 g after generation 38. It can be concluded that only a small fraction of the mutational variance could have been contributed by mutations with minor effects; if these had made a large contribution to the variance, a slow buildup in genetic variance would have been expected and would have contributed to an accelerating response in the later generations (Hill 1982b), the period when the plateau was most evident. A similar conclusion regarding a small contribution to the mutation pressure for fitness from minor mutations has been reached from analysis of data on fertility and viability in this experiment (Caballero and Keightley 1998).

I thank Bill Hill for continued interest and encouragement over the period of this experiment, Fiona Oliver for technical assistance,

Sara Knott for helpful advice on the analysis, Philippe Baret and two anonymous reviewers for helpful comments on the manuscript, and the Biotechnology and Biological Sciences Research Council and the Royal Society for support.

LITERATURE CITED

Bradford, G. E., and T. R. Famula, 1984 Evidence for a major gene for rapid postweaning growth in mice. Genet. Res. **44:** 293–308.

Caballero, A., and P. D. Keightley, 1998 Inferences on genome-wide deleterious mutation rates in inbred populations of Drosophila and mice. Genetica (in press).

Caballero, A., M. A. Toro and C. Lopez-Fanjul, 1991 The response to artificial selection from new mutations in *Drosophila melanogaster*. Genetics **127:** 89–102.

Caballero, A., P. D. Keightley and W. G. Hill, 1995 Accumulation of mutations affecting body weight in inbred mouse lines. Genet. Res. **65:** 145–149.

Clayton, G., and A. Robertson, 1955 Mutation and quantitative variation. Am. Nat. **89:** 151–158.

Cockerham, C. C., 1986 Modifications in estimating the number of genes for a quantitative character. Genetics **114:** 659–664.

Dietrich, W., H. Katz, S. E. Lincoln, H. Shin, J. Friedman *et al.*, 1992 A genetic map of the mouse suitable for typing intraspecific crosses. Genetics **131:** 423–447.

Elston, R. C., 1979 Major locus analysis for quantitative traits. Am. J. Hum. Genet. **31:** 655–661.

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, Ed. 4, Longman Scientific and Technical, Harlow, Essex, UK.

Festing, M., 1973 A multivariate analysis of subline divergence in the shape of the mandible in C57BL/Gr mice. Genet. Res. **21:** 121–132.

Festing, M. F. W., 1989 Inbred strains of mice, pp. 636–648 in *Genetic Variants and Strains of the Laboratory Mouse*, edited by M. F. Lyon and A. G. Searle. Oxford University Press, New York.

Fry, J. D., K. A. Deronde and T. F. C. Mackay, 1995 Polygenic mutation in *Drosophila melanogaster*: genetic analysis of selection lines. Genetics **139:** 1293–1307.

Genstat 5 Committee, 1993 *Genstat 5 Release 3 Reference Manual*. Clarendon Press, Oxford.

Go, R. C. P., R. C. Elston and E. B. Kaplan, 1978 Efficiency and robustness of pedigree segregation analysis. Am. J. Hum. Genet. **30:** 28–37.

Haley, C. S., G. J. Lee, R. Webb and S. A. Knott, 1993 Evidence on the genetic control of LH release in response to GnRH from crosses between selected lines of sheep. Livest. Prod. Sci. **37:** 153–167.

Heath, S. C., G. Bulfield, R. Thompson and P. D. Keightley, 1995 Rates of change of genetic parameters of body weight in selected mouse lines. Genet. Res. **66:** 19–25.

Hill, W. G., 1982a Rates of change in quantitative traits from fixation of new mutations. Proc. Nat. Acad. Sci. USA **79:** 142–145.

Hill, W. G., 1982b Predictions of response to artificial selection from new mutations. Genet. Res. **40:** 255–278.

Houle, D., B. Morikawa and M. Lynch, 1996 Comparing mutational variabilities. Genetics **143:** 1467–1483.

Juga, J., and R. Thompson, 1989 Estimation of variance components in populations selected over multiple generations. Acta Agric. Scand. **39:** 78–89.

Keightley, P. D., and S. Hawkins, 1991 Rimy: a new mutation affecting coat colour and body size. Mouse Genome **89:** 410.

Keightley, P. D., and W. G. Hill, 1992 Quantitative genetic variation in body size of mice from new mutations. Genetics **131:** 693–700.

Keightley, P. D., and G. Bulfield, 1993 Detection of quantitative trait loci from frequency changes at marker loci under selection. Genet. Res. **62:** 195–203.

Kimura, M., and J. F. Crow, 1963 On the maximum avoidance of inbreeding. Genet. Res. **4:** 399–415.

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

Lande, R., 1981 The minimum number of genes contributing to quantitative variation between and within populations. Genetics **99:** 541–553.

Lopez, M. A., and C. Lopez-Fanjul, 1993a Spontaneous mutation for a quantitative trait in *Drosophila melanogaster*. I. Response to artificial selection. Genet. Res. **61:** 107–116.

Lopez, M. A., and C. Lopez-Fanjul, 1993b Spontaneous mutation for a quantitative trait in *Drosophila melanogaster*. II. Distribution of mutant effects on the trait and fitness. Genet. Res. **61:** 117–126.

Lynch, M., and W. G. Hill, 1986 Phenotypic evolution by neutral mutation. Evolution **40:** 915–935.

Lyon, M., and A. Searle, 1989 *Genetic Variants and Strains of the Laboratory Mouse*, Ed. 2, Oxford University Press, Oxford.

Mackay, T. F. C., J. D. Fry, R. F. Lyman and S. V. Nuzhdin, 1994 Polygenic mutation in *Drosophila melanogaster*: estimates from response to selection in inbred strains. Genetics **136:** 937–951.

Merchante, M., A. Caballero and C. Lopez-Fanjul, 1995 Response to selection from new mutation and effective size of partially inbred populations. II. Experiments with *Drosophila melanogaster*. Genet. Res. **66:** 227–240.

Meyer, K., 1989 Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free approach. Genet. Sel. Evol. **21:** 317–340.

Nelder, J. A., and R. Mead, 1965 A simplex method for function minimization. Comput. J. **7:** 308–313.

Nuzhdin, S. V., J. D. Fry and T. F. C. Mackay, 1995 Polygenic mutation in *Drosophila melanogaster*—the causal relationship of bristle number to fitness. Genetics **139:** 861–872.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, 1992 *Numerical Recipes in C*, Ed. 2, Cambridge University Press, Cambridge.

Roberts, R. C., 1966 The limits to artiticial selection for body weight in the mouse. II. The genetic nature of the limits. Genet. Res. **8:** 361–375.

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

Wray, N. R., 1990 Accounting for mutation effects in the additive genetic variance-covariance matrix and its inverse. Biometrics **46:** 177–186.

Zeng, Z.-B., 1992 Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. Genetics **131:** 987–1001.

Zeng, Z.-B., D. Houle and C. C. Cockerham, 1990 How informative is Wright's estimator of the number of genes affecting a quantitative character? Genetics **126:** 235–247.

Communicating editor: Z-B. Zeng