

# Coat colour in mouse populations selected for weight gain: support for hitchhiking, not pleiotropy<sup>†</sup>

PHILIP W. HEDRICK\*

School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

(Received 5 September 2012; revised 29 November 2012; accepted 16 December 2012; first published online 24 January 2013)

## Summary

With many molecular markers in many species, research efforts in quantitative genetics have focused on dissecting these traits and understanding the importance of factors such as correlated response due to hitchhiking or pleiotropy. Here, in an examination of long-term selection experiments in mice, the evidence strongly supports the primary importance of hitchhiking on the coat colour loci *brown* and *dilute* in mice selected for high weight gain. First, the amount of observed change in coat colour allele frequency could not be explained by genetic drift alone, implying that selection was of high importance. Second, the allele frequency changes included reversals in the direction change, but there were still positive correlations in the early generations with differences in weight gain between the phenotypes. Third, the correlation between the change in allele frequencies and phenotypic difference in weight gain declined over time, consistent with the decay expected from linkage associations. Fourth, the changes at both loci in a short-term selection experiment for low weight gain were in the opposite direction than the changes in the contemporaneous related population selected for high weight gain.

## 1. Introduction

Determining the genetic basis of complex traits is the focus of extensive recent research using molecular genomic markers (Flint & Mackay, 2009). Before genomic examination was possible, long-term selection experiments and related crosses were used to characterize and estimate the amount of genetic response and the types of genetic variation important for quantitative traits (Comstock, 1996; Falconer & Mackay, 1996; Lynch & Walsh, 1998). These experiments statistically evaluated the importance of pleiotropy, epistasis, dominance and linkage as well as the genetic architecture, the number of genes and their effects influencing quantitative traits. For example, the role of various causes of the ultimate cessation of the response to directional selection, such as loss of genetic variation, negative genetic correlation with fitness traits and overdominance (heterozygote advantage), was of particular interest.

When there was polymorphism for visible traits in long-term selection experiments, potential insight into the relative significance of hitchhiking (linked gene effects) and pleiotropy on correlated response was possible. In general, associations between characters can be ephemeral, as that caused by linkage disequilibrium between genes influencing traits, or more persistent as that resulting from pleiotropic effects. Given selection for a particular trait, if the association with another trait is due to pleiotropy, then the response should be more constant, predictable and long term than if the association results from linkage associations which could decay and change over time.

The visible traits that were segregating in the mouse populations examined here were the coat colour genes *brown* (with recessive allele *b*) and *dilute* (with recessive allele *d*). The *brown* (*b*) gene is now known as *Tryp1* or Tyrosinase-related protein and codes for a melanosomal protein on chromosome 4 in the mouse (Bennett & Lamoreux, 2003). The homologous gene in humans is known as *TRYPI*, is located on chromosome 9, and variants can cause oculocutaneous albinism type 3 (OCA3). Interestingly, blond hair in Melanesians and colour polymorphism in a free-living population of sheep are caused by variants at the

\* Corresponding author: School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA. e-mail: philip.hedrick@asu.edu

<sup>†</sup> This manuscript is dedicated to the memory of Ralph Comstock, an exceptional researcher, teacher and colleague.

*TRYPI* locus (Gratten *et al.*, 2007; Kenny *et al.*, 2012). The *dilute* (*d*) gene is now known as *Myo5a* or Myosin Va and codes for the melanosome transport molecule on chromosome 9 in the mouse (Bennett & Lamoreux, 2003). The homologous gene in humans is known as MY05A, is located on chromosome 15, and variants can cause Griselli syndrome. Interestingly, one of the first inbred mouse strains was called DBA, named because it was homozygous at the coat colour mutations *dilute* (*d*), *brown* (*b*) and *agouti* (*a*) (Steingrimsdóttir *et al.*, 2006). C. C. Little began inbreeding this line in 1909 and it has been widely used in many research areas including cardiovascular biology, neurobiology, sensorineural research and quantitative genetics research discussed below.

In early studies, mice homozygous for both the recessive mutant brown and dilute phenotypes were observed to have larger body size than wild-type mice (Feldman, 1935; Green, 1935; Castle *et al.*, 1936; Castle, 1941). This effect could be attributed to pleiotropy where the coat colour genes directly influenced body weight or to linkage associations where the coat colour genes are associated by linkage to genes (quantitative trait loci (QTLs)) affecting body weight. For example, Castle *et al.* (1936) measured body weight for different backcrosses for both brown (*bb*) and dilute (*dd*) and their black (*Bb*) and non-dilute (*Dd*) sibs and found that larger size was associated with mice having the recessive phenotypes, implying a pleiotropic effect. Subsequently, MacArthur (1949) carried out a two-way selection experiment for body weight and found that his high line was mostly *bbdd* and his low line mostly *BBDD*, consistent with the findings of Castle *et al.* (1936).

On the other hand, Green (1935) measured body size in brown and heterozygous black mice and found that body size differences between colour phenotypes varied between different backcrosses and suggested that these effects were caused by different linkages between the *brown* locus and genes affecting body weight. Butler (1954) crossed MacArthur's high line with two different inbred lines homozygous for the non-dilute allele and found that non-dilute mice were larger in the  $F_2$  and backcrosses. Based on these results and other related ones, Butler (1954) suggested that differential linkage between the *dilute* locus and genes affecting body size was the source of these effects.

Ralph Comstock and his colleagues and students (Rahnefeld *et al.*, 1963) carried out long-term selection experiments for post-weaning weight gain (18–42 days of age) in mice from 1957 to 1976. These experiments were funded by the National Science Foundation at the University of Minnesota to determine with statistical tools the type of gene action, such as dominance and epistasis, the basis for plateaued response to selection and the number of

genes influencing weight gain in mice. In addition, the selected populations were segregating for these two coat colour loci (marker loci). In an analysis of the first part of these experiments (Hedrick & Comstock, 1968), it appeared that linkage associations between the coat colour genes and genes influencing weight gain were significant and that pleiotropy for weight gain by the coat colour genes was less likely.

Keightly & Bulfield (1993) examined populations selected for large or small body size which were started from crosses of inbred lines that differed at the *brown* and *dilute* loci (one of the lines they used was a DBA strain). The frequency of the *b* allele increased over 20 generations in the six large select lines (0.65) and decreased in the six small select lines (0.09), while the frequency of the *d* allele declined in the large select lines (0.15) and increased in the small selection lines (0.84). Keightly & Bulfield (1993) and Heath (1995) concluded that these effects were the result of linkage of the coat colour alleles to body size QTLs rather than a direct pleiotropic effect of the coat colour alleles on body size. Heath (1995) analysed these data and estimated that there was a QTL 10 cM from *dilute* with an effect on body weight of 0.40 g. On the other hand, Heath (1995) estimated that there was a QTL 45 cM from *brown* with an effect on body weight of 1.70 g, but that it was 'difficult to draw any firm conclusions ... about the QTL linked to *brown*.' Finally, from examination of 927  $F_2$  individuals, Morris *et al.* (1999) identified QTLs for body weight on chromosomes 4 and 9, consistent with previous findings of linkage of QTLs with *brown* and *dilute* genes (on chromosomes 4 and 9).

Recently, there has been renewed interest in the extent of pleiotropy for quantitative traits. For example, Stearns (2010) reviewed the history of the concept of pleiotropy and provided a prospective for pleiotropy in genomics today. Subsequently, Wagner & Zhang (2011) provided a theoretical context for pleiotropy for quantitative traits and suggested that universal pleiotropy (every gene influences every trait) was not empirically supported. In contrast, Hill & Zhang (2012) stated that the evaluation of the nature of pleiotropy is highly dependent upon statistical assumptions and that the extent of pleiotropy is not necessarily limited. In addition, genome-wide studies for complex diseases and traits in humans have found abundant evidence of pleiotropy (e.g. Sivakumaran *et al.*, 2011). More specifically, considering pleiotropic effects of coat colour genes on weight in mice, there is evidence that the gene *agouti* influences obesity (Klebig *et al.*, 1995) and that *Mc4r* (a G protein-coupled receptor related to the colour gene *Mc1r*) also has impacts on obesity (Ste. Marie *et al.*, 2000).

Here, data are examined from the remainder of the generations of the selection experiments carried out by R. Comstock and his colleagues and assessed

whether the data from the total experiments are consistent with an explanation based mainly on linkage associations of the coat colour variants with loci influencing weight gain or on pleiotropy. In addition, the potential role of genetic drift in the variation of allele frequencies over time and in generating linkage disequilibrium between the coat colour loci is examined.

## 2. Methods

The populations of mice examined here originated at different times by crossing the same two unrelated inbred lines (Rahnefeld *et al.*, 1963). The two inbred lines had been maintained by full-sib mating for 25 or more generations and were fixed for the mutant *brown* and *dilute* alleles (genotype *bbdd*) and the black and non-dilute alleles (genotype *BBDD*), respectively. As a result, the initial allele frequency in the selection populations was 0.5 for the coat colour alleles and any other alleles that differed between the lines. Also, because the selected populations were the result of a cross between lines, there was the maximum linkage disequilibrium possible in the initial generation. The two inbred lines were similar in weight gain. The  $F_2$  and  $F_3$  generations of the selection lines were produced by mating random individuals within the  $F_1$  and  $F_2$  generations, respectively. Selection for a post-weaning growth rate began in the  $F_3$ , which was designated as generation 1 for each of the populations. The selection criterion was individual post-weaning weight gain from 18 to 42 days and the selected mice were randomly mated with the restriction that no brother–sister matings were permitted (some other efforts were made to minimize inbreeding within the lines; Rahnefeld *et al.*, 1963). The first selection line S for increased weight gain was started in 1957 and the second one S' was started 15 generations later in 1961. Two other similar short-term selection populations, S'' and S''', and one low selection population, were started at different times later, and will be discussed only briefly here.

The number of litters per generation in a selection population was ordinarily 40 or more with 20 or more males used with each male mated with two or more females. The numbers were generally greater than these values and never fell short of these values by more than a very few individuals or litters. The effective population size when there are different numbers of female and male parents can be estimated as

$$N_e = \frac{4N_f N_m}{N_f + N_m}, \quad (1)$$

where  $N_f$  and  $N_m$  are the numbers of female and male parents (Wright, 1931; Hedrick, 2011). Given that  $N_f=40$  and that  $N_m=20$ , then an estimate of

$N_e$  is approximately 53, or perhaps somewhat more if there were more parents and litters and if other efforts to minimize inbreeding increased  $N_e$ . For example, for 10 generations in the early part of the S population, the mean number of male parents was 21.4 and the mean number of female parents was 65.2, giving  $N_e=64.4$  for this period. On average for the S and S' populations, 331 and 347 progeny per generation were scored and measured. Therefore, as basic numbers for the calculations and examinations below, it was assumed that there were 60 effective parents that produced 300 progeny per generation. Deviations from the number of parents were examined in relevant situations to determine the impact of other values.

Mice coat colour in progeny of each generation was classified as either black (*B-D-*), silver-grey (*B-dd*), brown (*bbD-*), or dilute brown (*bbdd*). Different colours of mice can be seen in Silvers (1979) which is also online at <http://www.informatics.jax.org/wksilvers/>. The frequency of the recessive allele *b* in each generation was estimated as

$$p_b = \left( \frac{N_{bb}}{N} \right)^{1/2}, \quad (2a)$$

where  $N_{bb}$  and  $N$  are the numbers of *bb* homozygotes and the total number of individuals scored (the frequency of allele *B* was  $p_B$ ). The frequency of the recessive *d* allele in each generation was estimated as

$$q_d = \left( \frac{N_{dd}}{N} \right)^{1/2}, \quad (2b)$$

where  $N_{dd}$  is the numbers of *dd* homozygotes (the frequency of allele *D* was  $q_D$ ).

To examine whether genetic drift could result in the observed changes in allele frequency, simulations were carried out where in each generation a given number of diploid offspring was generated with phenotypes as indicated and the allele frequencies were estimated as described above. From these individuals, a given number of individuals were randomly selected as parents for the next generation. The change in the estimated allele frequency was calculated for each generation and the average absolute value of the allele frequency change was determined. Simulations were run for 70 generations to mimic the S population and only simulated populations that were polymorphic for 70 generations were used. In all cases, the replicate–generation numbers used to calculate the mean average absolute allele frequency change were at least 500 000.

A general measure of linkage disequilibrium between two loci (here for the two coat colour loci) is

$$D = x_{bd} - p_b q_d. \quad (3a)$$

Table 1. Symbols indicating the numbers of the four phenotypic classes for the two coat colour loci and the marginal numbers for each locus

	<i>B-</i>	<i>bb</i>	
<i>D-</i>	$N_{B-D-}$	$N_{bbD-}$	$N_{D-}$
<i>dd</i>	$N_{B-dd}$	$N_{bbdd}$	$N_{dd}$
Total	$N_{B-}$	$N_{bb}$	Total

where  $x_{bd}$  is the frequency of gamete *bd* and is estimated as

$$x_{bd} = \left( \frac{N_{bbdd}}{N} \right)^{1/2}.$$

Linkage disequilibrium was estimated using the standardized measure of Lewontin (1964)

$$D' = \frac{D}{D_{\max}}, \quad (3b)$$

where  $D_{\max}$  is the maximum  $D$  possible for a given set of allele frequencies at the two loci. That is,  $D_{\max}$  is equal either to the lesser of  $p_Bq_d$  or  $p_bq_D$  if  $D$  is positive or to the lesser of  $p_Bq_D$  or  $p_bq_d$  if  $D$  is negative. Linkage disequilibrium significance was tested with the likelihood ratio statistic

$$Q = \frac{4ND^2}{p_B(2-p_B)q_D(2-q_D)}, \quad (4)$$

which is approximately  $\chi^2$  distributed with one degree of freedom (Hill, 1974).

In addition, the probability using Fisher's exact test that the level of linkage disequilibrium as extreme as that observed was calculated (Lewontin, 1995; Weir, 1996). In this case, the number of phenotypes of the four classes, *B-D-*, *B-dd*, *bbD-* and *bbdd*, and the marginal numbers of the four phenotypic classes, *B-*, *bb*, *D-* and *dd*, are as given in Table 1. The probability of an array with these numbers of phenotypes, given the marginal numbers of phenotypes in Table 1, is

$$\Pr(N_{B-D-}, N_{B-dd}, N_{bbD-}, N_{bbdd} | N_{B-}, N_{D-}) = \frac{N_{B-}! N_{bb}! N_{D-}! N_{dd}!}{N_{B-D-}! N_{B-dd}! N_{bbD-}! N_{bbdd}! N!}. \quad (5)$$

All possible samples with the same marginal totals were generated and the probability and the amount of linkage disequilibrium for each sample calculated. The cumulative probability of observing a given array by chance is then calculated by summarizing the probabilities for all the samples with linkage disequilibrium greater than, or equal to, that in the observed sample.

Finally, simulations were run designed to mimic the populations with a given number of offspring in each

generation in which the linkage disequilibrium was measured and a given number of parents for the next generation. In these simulations, no linkage was assumed between the two loci and linkage disequilibrium was generated from genetic drift resulting from the limited number of parents in each generation. Simulations were run for 70 generations as for population S, only replicates that were polymorphic for all generations were used, and at least 5000 values of each combination of the number of offspring and parents were generated. The amount of linkage disequilibrium was estimated both using the phenotype frequencies in the offspring, as in the actual data, and using the gametic frequencies known from the simulations.

To examine the effect of artificial selection for weight gain on allele frequency change in the coat colour loci, first the average weight gain of mice with the recessive phenotype (*bb* or *dd*) minus the average weight gain of mice with the dominant phenotype (*B-* or *D-*) was calculated. Second, the estimated change in allele frequency for the recessive allele between the generation in which the weight gain was measured and the next generation was calculated and then the correlation coefficient between these two values was calculated. For example, a positive correlation coefficient indicates that overall a higher weight gain of the recessive phenotype results in an increase in the recessive allele and that a lower weight gain of the recessive phenotype results in a decrease in the recessive allele.

### 3. Results

As a background to illustrate the effectiveness of the selection regimes, Fig. 1 gives the weight gain for populations S and S' over the course of the two experiments. Population S reached a plateau in generations 47–60 with an average weight gain of 20.4 g, nearly double the initial weight gain of 11.0 g (Comstock & Enfield, 1981). In other words, even though the two lines crossed to form the initial population had similar low weight gains, selection resulted in a large increase in weight gain, implying considerable genetic variation in the initial population. Although populations S and S' were started 15 generations apart (nearly 4 years), the responses for the shared generations were similar. For example, the mean weight gain for the last ten generations (36–45) of population S' was 20.4, the same average as that for the plateau level observed in the S population.

As further background information for these experiments, the estimate of heritability of weight gain in the narrow sense over the first 17 generations of selection for line S was about 0.25 for both males and females (Rahnefeld *et al.*, 1963). Using the selective response for weight gain estimated during the same

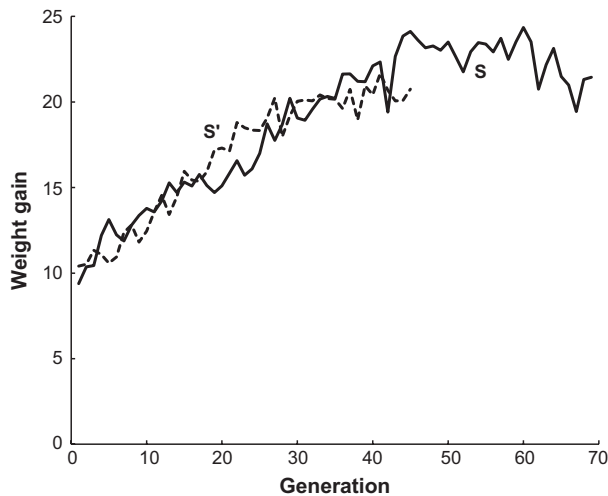


Fig. 1. The mean post-weaning weight gain from 18 to 42 days for mice in two long-term selection populations, S and S', for high weight gain.

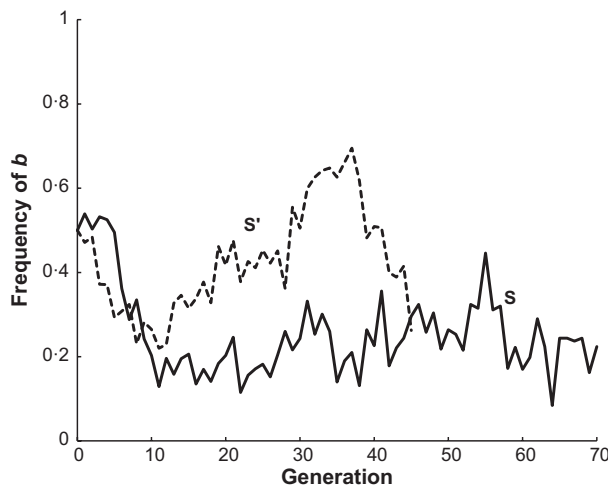


Fig. 2. The estimated frequency of the *brown* (*b*) allele for mice in two long-term selection populations, S and S', for high weight gain.

period and a theoretical approach, the estimate of the number of genes contributing to the response, assuming multiplicative gene action and no epistasis, was large and greater than 100 (Comstock & Enfield, 1981).

The estimated frequency of the *b* allele at the *brown* locus over the 70 generations of line S and the 45 generations of line S' are given in Fig. 2. Both populations initially declined in frequency from 0.5 to around 0.2 in generation 10. The frequency of *b* in S stayed at around this level for the remaining 60 generations. On the other hand, the frequency of *b* in S' increased to nearly 0.7 in generation 36 and then declined to 0.26 in the last generation 45. In other words, S' appeared to have two reversals in allele frequency from the initial decline while S had none. Neither population appeared to be approaching

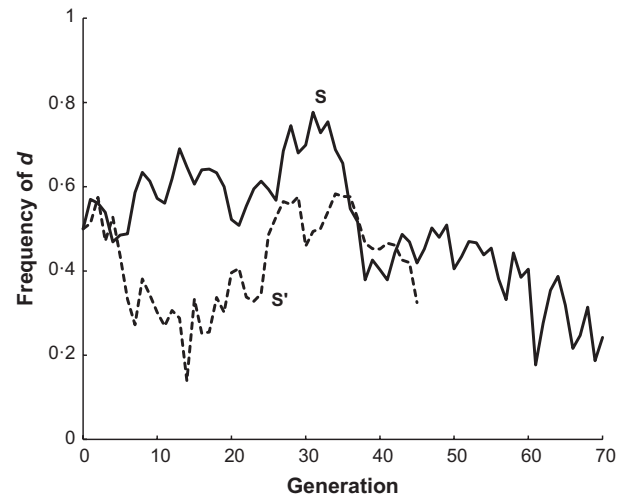


Fig. 3. The estimated frequency of the *dilute* (*d*) allele for mice in two long-term selection populations, S and S', for high weight gain.

fixation at the termination of the experiment. There are also data from two other short-generation populations, S'' and S''', not given in Fig. 2. In generations 1–3 of S'', the frequency of *b* was close to 0.5 and then in generations 4–6 increased to around 0.6. In generations 1–3 of S''', the frequency of *b* was close 0.5 but in generations 4 and 5 declined to nearly 0.3.

The estimated frequency of the *d* allele at the *dilute* locus is given for lines S and S' in Fig. 3. There is little change in both populations in the first few generations but then *d* increases in S to 0.78 in generation 31 and then declines to around 0.2 at the end of the experiment. On the other hand, S' initially declines to below 0.3, then increases to nearly 0.6, and then declines to nearly 0.3 at the end of the experiment. This pattern for *d* is similar to the two reversals in *b* allele frequency also observed in S'. Although both populations declined in frequency by the end of the experiments neither appeared to be approaching fixation. In generations 1 and 2 of S'', the frequency of *d* was close to 0.5 and then in generations 4–6 decreased to around 0.3. In generations 2–5 of S''', the frequency of *d* declined to 0.4.

Are these changes consistent with that expected by chance? Hedrick & Comstock (1968) demonstrated that in the first part of these experiments there were significant reversals in the direction of allele frequency change, suggesting that the early changes were the result of genes influencing weight gain linked to the coat colour genes. Here, let us examine if genetic drift generated by the limited number of parents could generate the changes observed over the total period of the populations.

The left side of Table 2 gives the observed average absolute change in allele frequency for populations S and S' for both loci. Overall, these observed values do not differ very much and the mean of the four values is

Table 2. The observed absolute average change in allele frequency for loci *b* and *d* in populations *S* and *S'* ( $\Delta p_b$  or  $\Delta q_d$ ) and the simulated absolute average change in allele frequency ( $\Delta p$ ) for different effective population sizes (number of parents)

Population	Locus	$\Delta p_b$ or $\Delta q_d$	$N_e$	$\Delta p$
<i>S</i>	<i>b</i>	0.058	40	0.044
	<i>d</i>	0.051	50	0.040
<i>S'</i>	<i>b</i>	0.053	60	0.037
	<i>d</i>	0.052	70	0.035
Mean		0.053		

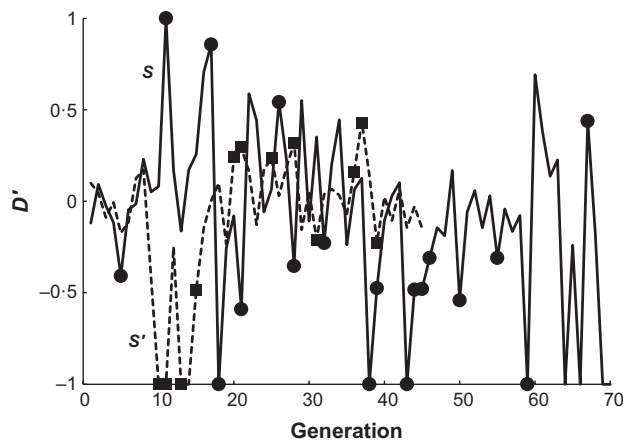


Fig. 4. The estimated amount of linkage disequilibrium between the *brown* (*b*) and *dilute* (*d*) alleles using the measure  $D'$  of Lewontin (1964) where the generations with significant linkage disequilibrium are indicated by solid circles for population *S* and solid squares for population *S'*.

0.053. Simulations were done to determine if this level of average change could be generated by a limited number of parents (small effective population size). The right side of Table 2 gives these average values for four different numbers of parents. All of these theoretically generated values are less than that observed. From the way each generation of the populations were set up with 20 males mated to two or more females, and efforts made to reduce inbreeding, it is unlikely that the effective population size is below 50. However, even with  $N_e = 40$ , the amount of change per generation (0.044) is below that observed (0.053). In other words, it seems unlikely that all of the variation observed in allele frequency over time was generated by genetic drift and suggests that selection was important causing changes in allele frequency.

The estimated linkage disequilibrium as measured by  $D'$  between the unlinked *b* and *d* loci is given in Fig. 4. Generations where there is significant disequilibrium using the likelihood ratio statistic between the loci are indicated by either solid circles (*S*) or solid squares (*S'*). There were very similar percentages of

generations with significant disequilibrium for the *S* population (18 of 70 or 26%) and for the *S'* population (12 of 45 generations or 27%). For both populations, the percentage of significant values greatly exceeded the 5% significance level expected.

$D' < 0$  generally occurred when there was a deficiency of *bbdd* individuals compared with linkage equilibrium expectations and sometimes no *bbdd* individuals were present which gives  $D' = -1$ .  $D' > 0$  values generally occurred when there was an excess of *bbdd* individuals compared with expectations or in a few cases where there was a deficiency of *bbD*- individuals. For population *S*, 14 of the 18 significant linkage disequilibrium values had  $D' < 0$  and most of these had either 0 or 1 *bbdd* individuals. Population *S'* had equal numbers of significant  $D' < 0$  and  $D' > 0$  values. Most of the  $D' < 0$  values occurred in the early generations when *bbdd* individuals were missing and the  $D' > 0$  values occurred in later generations from an excess of *bbdd* individuals over expectations.

The generations in populations *S* and *S'* for which there was significant linkage disequilibrium using the likelihood ratio statistic and the exact probability of observing this level of linkage disequilibrium are given in Appendix Tables A1 and A2 along with some other data about these generations. For both populations, the generations with exact probabilities less than 0.05 were subsets of the generations significant using the likelihood ratio statistic. For population *S*, there were 10 generations (14.3%) that had exact probabilities less than 0.05. That is, 8 of the generations that had significant linkage disequilibrium using the likelihood ratio statistic had exact probabilities greater than 0.05. For example, generation 11 had  $Q = 6.86$  but  $\text{Pr} = 0.102$ . In this case,  $D' = 1$  because no *bbD*- were observed while 4.1 were expected. Generation 18 had  $Q = 10.93$  but  $\text{Pr} = 0.126$ . In this case,  $D' = -1$  because no *bbdd* were observed while 1.8 were expected. For population *S'*, there were 8 generations (17.8%) that had exact probabilities less than 0.05 (four of the generations with significant linkage disequilibrium with the likelihood ratio statistic had exact probabilities greater than 0.05). For all generations, the mean observed value of the likelihood ratio statistic was 2.82 and 3.01 for populations *S* and *S'*, respectively.

In order to examine the potential impact of the method of linkage disequilibrium estimation on increasing the observed level of linkage disequilibrium, simulation results using the phenotypic approach above was compared with that estimated from the known simulated gametic frequencies (remember gametic numbers are not available from the actual populations but only from these simulations). Table 3 gives these estimates for three different numbers of parents. For 60 parents, the minimum expected number, the phenotypic and gametic approaches gave

Table 3. Simulation results with 300 offspring per generation and different numbers of parents for the amount of linkage disequilibrium estimated by  $Q$  and the proportion of significant  $Q$  values ( $> 3.84$ ) using the number of different phenotypes, as in the experiment, and using the number of gametes in the offspring

Number of parents	$Q$		Proportion $Q > 3.84$	
	Phenotype	Gamete	Phenotype	Gamete
60	1.48	1.44	0.103	0.102
80	1.39	1.07	0.092	0.062
100	1.29	0.86	0.083	0.040

very similar average  $Q$  values (1.44 and 1.48) and percentage of significant  $Q$  values (10.3 and 10.2%). Note that the proportion of significant  $Q$  values is greater than the 0.05 level expected, but less than that observed in the two populations, even when using the exact probabilities. Note also that the mean value of the likelihood ratio statistic from the simulations for 60 parents was much less than the mean observed value of the statistic.

By looking at the effect of different numbers of parents, we can determine the impact of genetic drift on the level of linkage disequilibrium. For 80 and 100 effective parents, the level of  $Q$  and the proportion of significant  $Q$  values was less than that for 60 parents. For example, when there were 100 parents using the number of phenotypes, the percentage of significant  $Q$  values declined from 10.3% for 60 parents to 8.3% for 100, indicating that genetic drift does generate some of the observed linkage disequilibrium.

To examine the potential effect of artificial selection changing the frequency of the two coat colour loci, the correlation between the weight gain difference between recessive and dominant phenotypes and the change in recessive allele frequencies is given in Table 4. The first period in populations S (generation 1–33) and S' (generations 1–18) are from Hedrick & Comstock (1968) where they observed a strong positive correlation for both loci in both populations and the correlation for the  $b$  locus in population S and for the  $d$  locus in population S' were significant even though there were reversals in allele frequencies. However, in the second part of the experiments examined here, generations 34–69 in population S and generations 19–42 in population S', the correlations in all four locus-population combinations were lower and none were statistically significant. Over all the generations, the correlations were non-significant for both loci for population S and for locus  $b$  in population S' while the correlation was still significant for locus  $d$  in population S', presumably because of the strong effect of the early generations.

Table 4. Correlation between the change in allele frequency and the difference in weight gain between dominant and recessive phenotypes for the brown and dilute loci observed in Hedrick & Comstock (1968) (generations 1–33 in populations S and generations 1–18 in population S') and for the rest of the generations examined here

Population	Generation	$b$	$d$
S	1–33	0.381**	0.272
	34–69	0.190	–0.065
	1–69	0.104	0.036
S'	1–18	0.367	0.624**
	19–44	0.169	0.175
	1–44	0.191	0.436**

\*\*  $P < 0.01$ .

A population selected for low weight gain was initiated from the S population in generation 8 for six generations. The pattern of allele frequency change for the coat colour loci in this population was different from the simultaneous S population selected for high weight gain (Table 5). In fact in these generations, the  $b$  allele declined while in the population selected for low weight gain it increased (the average difference in frequency between these lines was 0.123). A similar, but not as extreme, pattern occurred for the  $d$  allele with higher frequencies in the high population line and lower frequencies in the low selected population. In other words, the changes in the populations selected in the opposite direction for weight gain were also opposite in the change in coat colour allele frequency supporting the effect of selection on weight gain on the frequency of the coat colour alleles.

#### 4. Discussion

There is a history of experiments that have shown an association of coat colour alleles at the *brown* and *dilute* genes with body weight or weight gain. Overall, there is strong evidence that the patterns of change in populations segregating for these variants and the phenotypic values in experimental crosses are influenced by selective differences. Some early studies suggested that this association was the result of pleiotropic effects of the mutants on body weight but more recent examination suggests that QTLs with linked associations are more consistent with the data although it is not possible to completely exclude pleiotropic effects. Most of these studies used crosses between inbred or differentiated lines to initiate their experiments so that high linkage disequilibrium between the coat colour mutants and alleles at any genes (QTLs) that were different between the parental lines was initially expected.

Table 5. The observed frequencies of alleles *b* and *d* in contemporaneous generations of population *S* selected for high weight gain and a population selected for low weight gain originated from generation 7 of population *S*

Generation	<i>b</i>			<i>d</i>		
	High (S)	Low	High–Low	High (S)	Low	High–Low
8	0.366	0.314	0.052	0.636	0.574	0.062
9	0.242	0.316	–0.074	0.606	0.688	–0.082
10	0.203	0.313	–0.110	0.580	0.567	0.013
11	0.129	0.480	–0.351	0.556	0.480	0.076
12	0.201	0.255	–0.054	0.592	0.500	0.092
13	0.158	0.358	–0.200	0.687	0.580	0.107
Mean	0.216	0.339	–0.123	0.610	0.565	0.045

In experiments examined over time, the pattern of change in allele frequency for the coat colour alleles varied between replicate populations, over time within the populations and over different studies. For example, although both populations *S* and *S'* examined here selected for high weight gain initially had reduced frequency of the *b* allele, the pattern of change after the first 10 generations was different with the *S'* population increasing greatly in allele frequency and *S* having a low frequency throughout the experiment. The two populations diverged at the *d* allele early in the experiment and had reversals in allele frequency at different times. These patterns would be unlikely if pleiotropy were the most important factor influencing allele frequency change. In the similar experiments examined by Keightly & Bulfield (1993) and Heath (1995), the *b* allele increased in frequency in populations selected for high body weight (and decreased in lines selected for low body weight) and the *d* allele declined in frequency in populations selected for high body weight (and increased in lines selected for low body weight). A likely explanation for the dissimilarities in the two experiments was that the inbred lines used to establish these populations were different so that different associations of linked QTLs with the mutant alleles occurred.

Several other findings point to the importance of selection at linked QTLs, relative to pleiotropy, influencing the change in coat colour allele frequency in populations selected for size. First, the amount of change in allele frequency found in populations *S* and *S'* was too great to be explained only by genetic drift, suggesting that selection was an important factor. Second, the association of allele frequency change with differences in weight gain was highest in the first part of the experiment with *S* and *S'* and was lower in the last part of the experiment, consistent with the expected decay in association for linked variants. On the other hand, if pleiotropy was important, then these associations would not be expected to decline over time and if there were pleiotropic differences, then fixation or loss of the mutants might be expected by the end of the long-term experiments. The average

frequency of *b* for populations *S* and *S'* was 0.243 and the average frequency for *d* was 0.284 at the end of the experiments, suggesting some selection against both mutants, but both coat colour mutants were still segregating at a substantial frequency. Finally, Heath (1995) found that the patterns observed in the Keightly & Bulfield (1993) crosses were consistent with linked QTLs for both the *b* and *d* alleles and not with pleiotropy (for pleiotropy, the QTLs would map to the same map location as the mutants), although the QTL linked to the *d* allele appeared more closely linked than the QTL linked to the *b* allele.

A significant difference between the experiments reported here and those examined by Keightly & Bulfield (1993) and Heath (1995) was that after 20 generations, four of their six low populations were fixed for *b*, two of their six high populations were fixed for *d* and two of their six low populations were fixed for the *D* allele. In other words, in 8 of 24 cases (33%) their populations had become fixed by generation 20. On the other hand, after 70 generations in population *S* and 45 generations in populations *S'* neither population appeared close to fixation for either the *b* or *d* alleles. Part of the difference may be the effective population size and genetic drift in the situations. Keightly & Bulfield (1993) estimated that the effective population size within their lines was about 23 while a reasonable general estimate for *S* and *S'* was around 60, several times larger. In addition, assuming that there were QTLs initially in strong linkage disequilibrium, and then their impact on allele frequency change would be greatest in the early generations before the linkage disequilibrium had decayed and where genetic drift may have augmented selective changes.

The observed number of generations with significant linkage disequilibrium was much higher than the expectation. To provide an evaluation of this, several potential factors were examined using simulations that might have contributed. First, some of the high proportion of generations with significant linkage disequilibrium is reduced if exact probabilities are used instead of the likelihood ratio statistic. Second,



the method of estimating using phenotype numbers instead of gametic numbers (not available from the experimental populations) has little effect on the proportion of significant linkage disequilibrium values, except when the number of parents is high. Finally, genetic drift because of a limited number of effective parents can increase the proportion of significant linkage disequilibrium values. Taking these factors into consideration, it does not seem necessary to invoke some further factors, such as selection, to generate what at first appeared to be a high level of linkage disequilibrium between the two loci.

Examination of the coat colour loci over these long-term experiments provides more support for the hypothesis that associations generated by linkage of

QTLs to the coat colour loci were more important than pleiotropic effects. Overall, the length and size of these experiments are generally unequalled today (with some exceptions, e.g., Renne *et al.* 2003; Dudley & Lambert, 2004), and the analysis of these data provides important insights into selection for quantitative traits.

Partial support for this research to P. W. H. was provided by the Ullman Professorship. The original experiments were supported by the National Science Foundation under grants to R. E. Comstock. I thank Elizabeth King, 'queen of the mouse data', for her conscientious and thorough efforts to summarize information from the original data books compiled by the students and technicians of Ralph Comstock. I also thank two anonymous reviewers for their comments on the manuscript.

## Appendix

Table A1. The generations in experiment *S* where the likelihood ratio statistic was significant ( $> 3.84$ ) and the exact probability was  $< 0.05$  (\* indicates generations where both were significant) along with allele frequencies, number of *bbdd* genotypes and *D'* in these generations

Generation	<i>b</i>	<i>d</i>	Obs. ( <i>bbdd</i> )	<i>D'</i>	<i>Q</i>	Exact Prob.
5	0.495	0.485	6	-0.408	19.81	0.000*
11	0.129	0.561	6	1.00	6.86	0.102
17	0.170	0.642	9	0.857	6.60	0.040*
18	0.141	0.633	0	-1.00	10.93	0.126
21	0.246	0.508	1	-0.590	11.89	0.008*
26	0.152	0.568	9	0.542	4.64	0.044*
28	0.260	0.745	5	-0.354	13.36	0.002*
32	0.253	0.728	6	-0.228	4.78	0.048*
38	0.131	0.379	0	-1.00	4.10	0.396
39	0.264	0.426	1	-0.475	4.29	0.106
43	0.221	0.487	0	-1.00	19.59	0.015*
44	0.244	0.469	1	-0.483	4.77	0.082
45	0.296	0.419	1	-0.479	4.51	0.089
46	0.324	0.452	4	-0.310	4.51	0.057
50	0.264	0.405	1	-0.541	7.16	0.033*
55	0.446	0.454	5	-0.310	6.38	0.023*
59	0.222	0.385	0	-1.00	5.11	0.318
67	0.237	0.247	4	0.439	5.91	0.003*

Table A2. The generations in experiment *S'* where the likelihood ratio statistic was significant ( $> 3.84$ ) and the exact probability was  $< 0.05$  (\* indicates generations where both were significant) along with allele frequencies, number of *bbdd* genotypes and *D'* in these generations

Generation	<i>b</i>	<i>D</i>	Obs. ( <i>bbdd</i> )	<i>D'</i>	<i>Q</i>	Exact Prob.
10	0.265	0.302	0	-1.00	8.64	0.138
11	0.220	0.269	0	-1.00	4.59	0.342
13	0.327	0.288	0	-1.00	14.61	0.037*
15	0.315	0.333	1	-0.485	4.42	0.092
20	0.417	0.396	20	0.244	7.40	0.001*
21	0.476	0.406	20	0.298	7.54	0.013*
25	0.453	0.486	17	0.237	4.48	0.056
28	0.362	0.558	21	0.319	5.73	0.023*
31	0.588	0.494	18	-0.209	4.51	0.048*
36	0.661	0.576	94	0.163	5.97	0.018*
37	0.695	0.525	91	0.434	24.95	0.000*
39	0.482	0.453	17	-0.225	9.38	0.004*

## References

- Bennett, D. C. & Lamoreux, M. L. (2003). The color loci of mice – a genetic century. *Pigment Cell Research* **16**, 333–344.
- Butler, L. (1954). The effect of the coat colour dilution gene on body size in the mouse. *Heredity* **8**, 275–278.
- Castle, W. E. (1941). Influences of certain colour mutations on body size in mice, rats and rabbits. *Genetics* **26**, 117–191.
- Castle, W. E., Gates, W. H., Reed, S. C. & Law, L. W. (1936). Studies of a size cross in mice. II. *Genetics* **26**, 310–323.
- Comstock, R. E. (1996). *Quantitative Genetics with Special Reference to Plant and Animal Breeding*. Ames, IA: Iowa State University Press.
- Comstock, R. E. & Enfield, F. D. (1981). Gene number estimation when multiplicative genetic effects are assumed – growth in flour beetles and mice. *Theoretical and Applied Genetics* **59**, 373–379.
- Dudley, J. W. & Lambert, R. J. (2004). 100 generations of selection for oil and protein in corn. *Plant Breeding Reviews* **24**, 79–100.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow, UK: Addison Wesley Longman.
- Feldman, H. W. (1935). The brown variation and growth of the house mouse. *American Naturalist* **69**, 370–374.
- Flint, J. & Mackay, T. F. C. (2009). Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research* **19**, 723–733.
- Gratten, J., Beraldi, D., Lowder, B. V., McRae, A. F., Visscher, P. M., Pemberton, J. M. & Slate, J. (2007). Compelling evidence that a single nucleotide substitution in *TYRP1* is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B* **274**, 619–626.
- Green, C. V. (1935). Apparent changes with age in X-over between colour and size genes in mice. *Journal of Genetics* **30**, 101–106.
- Heath, S. C. (1995). Inferences on the Genetic Control of Quantitative Traits from Selection Experiments. PhD. Thesis, University of Edinburgh.
- Hedrick, P. W. (2011). *Genetics of Populations*, 4th edn. Sudbury, MA: Jones and Bartlett Publishers.
- Hedrick, P. W. & Comstock, R. E. (1968). Role of linkage in gene frequency change on coat color alleles in mice. *Genetics* **58**, 297–303.
- Hill, W. G. (1974). Estimation of linkage disequilibrium in randomly mating populations. *Heredity* **33**, 229–239.
- Hill, W. G. & Zhang, X.-S. (2012). On the pleiotropic structure of the genotype-phenotype map and the evolvability of complex organisms. *Genetics* **190**, 1131–1137.
- Keightley, P. D. & Bulfield, G. (1993). Detection of quantitative trait loci from frequency changes of marker alleles under selection. *Genetical Research* **62**, 195–203.
- Kenny, E. E., Timpson, N. J., Sikora, M., Yee, M.-C., Moreno-Estrada, A., Eng, C., Huntsman, S. Burchard, E. G., Stoneking, M., Bustamante, C. D. & Myles, S. (2012). Melanesian blond hair is caused by an amino acid change in *TYRP1*. *Science* **336**, 554.
- Klebig, M. L., Wilkinson, J. E., Geisler, J. G. & Woychik, R. P. (1995). Ectopic expression of the agouti gene in transgenic mice causes obesity, features of type II diabetes, and yellow fur. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 4728–4732.
- Lewontin, R. C. (1964). The interaction of selection and linkage I. General considerations; heterotic models. *Genetics* **49**, 49–63.
- Lewontin, R. C. (1995). The detection of linkage equilibrium in molecular sequence data. *Genetics* **140**, 377–388.
- Lynch, M. & Walsh, R. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates.
- MacArthur, J. W. (1949). Selection for small and large body size in the house mouse. *Genetics* **34**, 194–209.
- Morris, K. H., Ishikawa, A. & Keightley, P. D. (1999). Quantitative trait loci for growth traits in C57BL/6J × DBA/2J mice. *Mammalian Genome* **10**, 225–228.
- Rahnefeld, G. W., Boylan, W. J., Comstock, R. E. & Singh, M. (1963). Mass selection for post-weaning growth in mice. *Genetics* **48**, 1567–1583.
- Renne, U., Langhammer, M., Wytrwat, E., Dietl, G. & Bünger, L. (2003). Genetic-statistical analysis of growth in selected and unselected mouse lines. *Journal of Experimental Animal Science* **42**, 218–232.
- Silvers, W. K. (1979). *The Coat Colors of Mice: A Model for Mammalian Gene Action and Interaction*. Berlin: Springer-Verlag (<http://www.informatics.jax.org/wksilvers/>, revised 2008).
- Sivakumaran, S., Agakov, F., Theodoratou, E., Prendergast, J. G., Zgaga, L., Manolio, T., Rudan, I., McKeigue, P., Wilson, J. F. & Campbell, H. (2011). Abundant pleiotropy in human complex disease and traits. *American Journal of Human Genetics* **89**, 607–618.
- Stearns, F. W. (2010). One hundred years of pleiotropy: a retrospective. *Genetics* **186**, 767–773.
- Ste Marie, L., Miura, G. I., Marsh, D. J., Yagloff, K. & Palmiter, R. D. (2000). A metabolic defect promotes obesity in mice lacking melanocortin-4 receptors. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 12339–12344.
- Steingrimsson, E., Copeland, N. G. & Jenkins, N. A. (2006). Mouse coat color mutations: from fancy mice to functional genomics. *Developmental Dynamics* **235**, 2401–2411.
- Wagner, G. P. & Zhang, J. (2011). The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics* **12**, 204–213.
- Weir, B. S. (1996). *Genetic Analysis II*. Sunderland, MA: Sinauer Associates.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.