CHANGES OF GENETIC VARIANCE AND FREQUENCIES OF MARKER ALLELES IN MOUSE LINES SELECTED ON 6 WEEK WEIGHT

P.D. Keightley[1], S.C. Heath[1], T. Hardge[2*], L. May[1], A.L. Archibald[2] and G. Bulfield[2]
[1]ICAPB, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JT Scotland; [2]Roslin Institute (Edinburgh), Midlothian EH25 9PS, Scotland, [*Present address: Inst. für Grundlagen der Nutztierwissenschaften, Fachgebiet Züchtungsbiologie und molekulare Tierzüchtung, Humboldt Universität zu Berlin, Invalidenstr. 42, Berlin, Germany.]

SUMMARY

Changes of genetic variance and frequencies at marker loci are described for an experiment on divergent selection on 6 week weight of mice which used a cross between two inbred strains as its base population. Estimates of genetic and environmental components of variation were obtained from the regression of response on within family selection differential and by animal model REML. Heritabilities estimated by both methods were about 0.2, and in close agreement. Despite the limited genetic base, there was substantial genetic variance. A model was fitted in which variance components were allowed to vary in different time periods and between sets of lines. This analysis suggested that the genetic variance increased substantially during the latter stages of the selection experiment. In several regions of the genome, marker alleles diverged in frequency between the high and low line replicates to a greater extent than would be expected by chance, suggesting the presence of QTL with effects of at least 0.3 phenotypic standard deviations.

INTRODUCTION

The use of molecular markers to detect linkage to quantitative trait loci (QTL) is becoming widespread in commercial and experimental species. Information is rapidly being accrued on the distributions and numbers of QTL, dominance effects at individual loci and interactions between loci, and their linkage relationships. The ultimate goal of many such studies is the isolation of some of the genetic factors influencing the quantitative traits.

This paper describes an experiment which was initiated with the aim of locating QTL influencing growth rate in mice by using artificial selection to generate changes in allele frequencies between replicates subjected to divergent selection. The base population for the selection lines (X lines) was the F_2 of two well-characterised inbreds, C57BL/6J and DBA/2J. By starting from an F_2, the initial state of the population was well defined, as gene frequencies were at 0.5. Frequencies of alleles at marker loci are expected to change as a consequence of linkage to QTL, which contribute to variation in the trait under selection. Only markers in tight linkage to QTL are expected to show substantial changes in frequency, however. It is possible to infer the magnitudes of effects of QTL from the gene frequency changes at the associated markers. This strategy for detecting QTL has similarities to 'selective genotyping' in *e.g.* an F_2 population (Lebowitz *et al.*, 1987), in which only the extremes of the phenotypic distribution are typed, but in the present case the selection is extended over many generations, and there is greater opportunity for recombination to break up

associations between markers and QTL. This paper describes changes of genetic variance of 6 week weight which occurred during the period of selection, and allele frequency changes at various classes of marker loci distributed throughout the genome after 21 generations of selection.

MATERIALS AND METHODS

Selection lines. The base population (generation 0) was the F_2 of the inbred lines C57BL/6J and DBA/2J. Within family selection on weight at six weeks was practised for 21 generations in 6 high and 6 low lines of 8 pairs of parents per generation. An unselected control line was also maintained. A mating scheme which minimises matings between close relatives was used (Falconer, 1973).

Analysis of quantitative data. Heritability (h^2) of 6 week weight was estimated from the regression of phenotypic mean on within family selection differential, and by animal model REML (AM-REML) with the programs of Meyer (1988). Changes in genetic and environmental components of variance were estimated using an AM in which variance components were allowed to vary across groups of lines and in different time periods (Beniwal *et al.*, 1991).

Molecular markers. After 21 generations of selection, phenotype frequencies were measured at 12 non-ecotropic retrovirus markers by Southern blotting (Blatt *et al.*, 1983; Keightley and Bulfield, 1993). These are dominant markers. Allele frequencies were measured at a total of 41 simple sequence length polymorphism (SSLP) markers (Research Genetics Inc.) which vary between the two inbreds (Love *et al.*, 1990; Cornall *et al.*, 1991; Dietrich *et al.*, 1992). SSLP's were amplified by PCR in 20μl reaction volumes in conditions that yielded sufficient product for visualisation by ethidium bromide staining after separation on 20cm polyacrylamide gel. Loci in which the products from the two alleles differed in length by more than 5% were chosen, in which cases the three genotypes could be easily distinguished.

RESULTS

Response to artificial selection and realised heritability. Mean 6 week weight for each of the six high and low replicates, the control, the parental strains, and the F_1 and F_2 are shown in Fig. 1. There was strong heterosis because the F_1 phenotypic mean was about 3g higher than the inbred parents (which only differ slightly for body size). This is evidence of directional dominance of genes affecting body size. The hybrid vigour appeared to break down only slightly in the F_2, but because the two groups were not bred at the same time, it is difficult to give precise relative values for the F_1 and F_2. The hybrid breakdown was perhaps greater between the F_2 and the F_3 (*i.e.* generations 0 and 1 of selection). The pattern of the selection response was different in the high and low line replicates. After the initial apparent reversal, the upward responses appeared to be almost linear, showing little sign of slowing down in later generations, and there was surprisingly little variation among the replicate means. The downward selection lines responded little between generations 1 and 5, but response accelerated between generations 7 and 12, and thereafter appeared to attenuate. One of the low replicates showed a consistently higher response after generation 8, for which plausible

explanations are a new mutation or rare recombination event. The gradual and slight decline in mean performance of the control line was presumably a consequence of inbreeding.

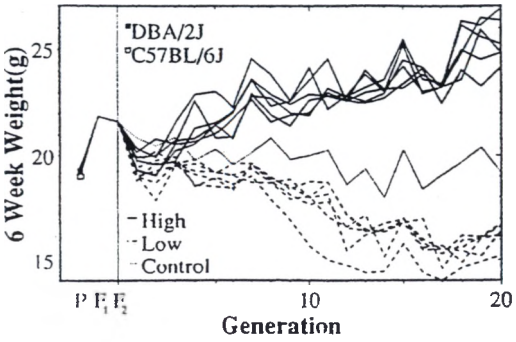

Fig. 1. Response to artificial selection on 6 week weight to generation 20 in the X lines.

Realised within family heritability estimated from the regression of response on within family selection differential was 0.22 with a standard error of 0.02 computed from the mean among pairs of replicates. In this case, within family heritability is almost equivalent to individual heritability. Although the base population was a cross between two inbred strains of mice, and these hardly differed in body size, substantial genetic variation was revealed by random segregation of chromosomes and by recombination, and yielded selection responses only marginally smaller than from selection experiments with more complex base populations (*e.g.* Macarthur, 1944; Falconer, 1973).

AM-REML analysis. Log transformed data from the selection experiment were analysed by AM-REML with genetic, litter and random environmental variances, and fixed effects of sex × line × generation, parity and litter size. The heritability estimate was 0.18, which is close to the realised heritability estimate. A more complex analysis in which separate variance components were fitted in different time periods and to high and low replicates separately revealed, however, substantial changes in genetic variance during the course of the experiment. The variance component estimates are summarised in Table 1. The genetic variance estimate increased substantially in the low line until the last 5 generations, when the estimate was lower. The estimate of the genetic variance in the high lines between generations 5 and 10 was very small, but the variance estimate increased thereafter. The changes of variance were mirrored in the selection responses, particularly for the low lines which which showed an accelerating pattern of response, and an attenuation around generation 15. The responses in the high line replicates were rather more linear, so the very small estimate of h^2 between generations 5 and 10 is somewhat surprising. The likelihood of models with different variance components fitted to different time periods and groups of lines is very much higher than that for homogeneous variances, however. Some of the increase in variance could have come from changes in dominance and interaction variance as allele frequencies change as a consequence of selection. Alternatively, the breakdown of repulsion combinations of tightly linked QTL could have

contributed.

	σ^2_A		h^2	
Generation	Low	High	Low	High
0-4	37	25	0.31	0.22
5-9	43	3	0.37	0.04
10-14	118	33	0.57	0.28
15-20	27	44	0.19	0.34

Table 1. AM-REML estimates of relative additive genetic variance (σ^2_A, arbitrary units) and h^2 from a model in which separate additive, litter and random environmental components were fitted to high and low lines in different time periods of the selection experiment.

Molecular markers. Using a probe which hybridises to a large class of non-ecotropic retroviruses, frequencies of 12 retrovirus marker loci were measured by Southern blotting (Keightley and Bulfield, 1993). Observations at a marker locus consist of phenotype allele frequencies for each independent replicate in the high and low lines. A method based on maximum likelihood was developed to estimate magnitudes of effects of QTL associated with the markers under a model of complete linkage between marker loci and QTL (Keightley and Bulfield, 1993). The method consists of computing the distribution of allele frequency in a finite population after t generations at a diallelic locus with initial gene frequency of 0.5, at which there is additive gene action with a difference in selective value between the homozygotes of s. The likelihood of an observed allele frequency at a marker is the density of the frequency distribution at that point. The overall likelihood for the complete data set is maximized as a function of s. It is then straightforward to convert the ML estimate of s to a scale of allelic effect on the trait. Because complete linkage is assumed, effects of QTL tend to underestimated. In the region of three of the retrovirus markers, divergences of allele frequency were large and significant at the 5% level. At retrovirus locus *Mpmv-2* on chromosome 11, the estimated mean frequency in the low lines was 0, and was 0.8 in the high lines. This large divergence in allele frequency from the starting frequency of 0.5 is explained by a QTL with a selective difference between the homozygotes of 0.3, which translates to an effect of about 0.5 phenotypic s.d.

With the development of SSLP markers for the mouse (Love *et al.*, 1990; Cornall *et al.*, 1991; Dietrich *et al.*, 1992), it is feasible to measure frequency changes at markers in the X lines in very great detail, but only worthwhile to do so at a resolution of c.10cM because marker allele frequency changes becoming increasingly positively correlated with decreasing map distance. Out of a total of more than 50 markers so far typed in the X lines, 41 are sufficiently far apart (*i.e.* > 10cM) as to be considered independent. The frequency distribution of ML estimates of effects in phenotypic s.d. units associated with 41 SSLP markers is shown in Fig. 2. A theoretical curve, computed by Monte Carlo simulation, shows the expected frequency distribution if markers have zero effect.

The observed distribution of effects associated with markers contains more weight in the tails than the theoretical distribution. The corresponding markers are therefore likely to be linked to QTL affecting 6 week weight.

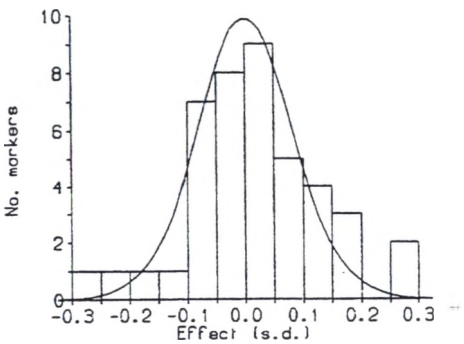

Fig. 2. Histogram showing the distribution of estimated effects in phenotypic s.d. units associated with 41 SSLP markers. The curve is the distribution expected if markers are not associated with QTL. The area under the curve and histogram are the same.

DISCUSSION

During the course of the selection experiment, the frequencies of alleles at two coat colour markers (*brown* and *dilute*) which differ between the two inbreds showed large and consistent divergences between the high and low line replicates. In regions of four other chromosomes (5, 11, 13 and 17), allele frequency divergences at SSLP markers were such that natural log likelihood ratios were greater than 4 (where the null hypothesis is neutral genetic drift). In the region of the *dilute* locus (chromosome 9) and on chromosome 11, several markers have been typed, and it has been found that the divergences of allele frequencies have occurred over a fairly large area (about 30cM). This suggests that there may be more than one gene involved in the response in these regions, but a statistical test for this has yet to be developed. On chromosome 17, however, there is a sharp peak in the change of marker allele frequency which suggests a single gene. It will be possible to include linkage in the analysis by combining allele frequencies at more than one marker in order to simultaneously estimate QTL effects and locations with similar maximum likelihood methods as described above. It may also be possible to compare likelihoods of models with, for example, one or several QTL in a region.

Under the assumption of homogeneous genetic and environmental components of variance, the base population heritability estimate was 0.18, which was close to the realised heritability estimate. If marker loci behave independently and QTL effects are additive and completely linked to markers, the heritability associated with a marker locus in the base population is $(\hat{a}^2 - V(\hat{a}))q_0(1 - q_0)/(2\sigma_P^2)$, where σ_P^2 is the phenotypic variance, q_0 is the initial gene frequency, $\hat{a}$ is the estimate

of the effect of the QTL on the trait, and $V(\hat{a})$ is the variance of the estimate. The estimate can be computed by maximum likelihood as described above, and its variance from the curvature of likelihood about the maximum. What proportion of the total genetic variance is associated with the markers typed in the X lines? The answer to this question is not straightforward because linked markers are non-independent. If, however, a subset of markers is chosen each of which is at least 10cM from all others, the total heritability associated with these is about 6%. It is likely therefore that further regions of large effect remain to be uncovered.

The pattern of the selection responses and increases of genetic variance revealed by the AM-REML analysis suggest that non-additive gene action may have made important contributions. The changes of allele frequencies at marker loci provide little information to distinguish between additive and non-additive effects, however. Changes of marker allele frequencies in regions that have been examined in detail do not so far suggest QTL closely linked in repulsion, which could also lead to an increase in genetic variance.

Acknowledgements. We thank W.G. Hill for helpful comments. This work was funded by the Agricultural and Food Research Council and the Royal Society. TH was supported by a bursary from the 'Leopoldina', Berlin.

REFERENCES

BENIWAL, B.K., HASTINGS, I.M., THOMPSON, R. and HILL, W.G. (1991) Heredity, 69: 361-371.

BLATT C., MILEHAM, K., HAAS, M, NESBITT, M.N., HARPER, M.E. and SIMON, M.I. (1983) P.N.A.S. 80: 6298-6302.

CORNALL, R.J., AITMAN, T.J, HEARNE, C.M. and TODD, J.A. (1991) Genomics, 10: 874-881.

DIETRICH, W., KATZ, H., LINCOLN, S.E., SHIN, H.-S., FRIEDMAN, J.,DRACOPOLI, N. and LANDER, E.S. (1992) Genetics, 131: 423-447.

FALCONER, D.S. (1973) Genet. Res., 22: 291-321.

KEIGHTLEY, P.D., and BULFIELD, G. (1993) Genet. Res., 66: 195-203.

LEBOWITZ, R.J., SOLLER, M. AND BECKMANN, J.S. (1987) Theor.Appl.Genet., 73: 556-562.

LOVE, J.M., KNIGHT, A.M., McALEER, M.A. and TODD, J.A. (1990) Nucleic Acids Res., 18: 4123-4130.

MACARTHUR, J.W. (1944) Amer. Nat., 78: 142-157.

MEYER, K. (1988) J. Dairy Sci., 71 (suppl 2): 33.