A Genetic Map of Quantitative Trait Loci for Body Weight in the Mouse

Peter D. Keightley,* Torsten Hardge,[†,1] Linda May*,[2] and Grahame Bulfield[†]

**Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland and [†]Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland*

Manuscript received April 6, 1995
Accepted for publication September 8, 1995

ABSTRACT

The genetic basis of body weight in the mouse was investigated by measuring frequency changes of microsatellite marker alleles in lines divergently selected for body weight from a base population of a cross between two inbred strains. In several regions of the genome, sharp peaks of frequency change at linked markers were detected, which suggested the presence of single genes of moderate effect, although in several other regions, significant frequency changes occurred over large portions of chromosomes. A method based on maximum likelihood was used to infer effects and map positions of quantitative trait loci (QTLs) based on genotype frequencies at one or more marker loci. Eleven QTLs with effects in the range 0.17–0.28 phenotypic standard deviations were detected; but under an additive model, these did not fully account for the observed selection response. Tests for the presence of more than one QTL in regions where there were large changes of marker allele frequency were mostly inconclusive.

BODY weight in mice is a typical polygenic trait that has been extensively studied as a model mammalian quantitative trait, but its genetic basis is poorly understood. The number, locations and effects of the individual genes contributing to natural variation in the trait are all unknown. Mutations of major effect at several loci alter some aspect of growth or lead to obese phenotypes (LYON and SEARLE 1989; FRIEDMAN and LEIBEL 1992), but it is unknown whether alleles of smaller effect at these loci contribute to quantitative genetic variation. The development of dense genetic maps based on molecular markers now provides opportunities to resolve quantitative genetic variation into individual regions of the genome influencing traits [quantitative trait loci (QTLs)]. This approach is not new (SAX 1923; THODAY 1961), but earlier studies depended on visible genetic markers which were limited in number and potentially affected the traits under investigation. Now, there are dense genetic maps in many species based on neutral genetic markers. The standard strategy for detection of QTLs is to cause the segregation of QTLs and marker loci in crosses where there is linkage disequilibrium between markers and QTLs (usually an F_2 or backcross population derived from inbred or widely differing strains) and to detect statistical associations between the trait value and marker genotypes. Information from two or more linked markers may be used simultaneously to estimate QTL effects and map positions (LANDER and BOTSTEIN 1989; HALEY and KNOTT 1992; JANSEN and STAM 1994; ZENG 1994). In plants and experimental animal species, it is also possible to infer QTL locations from segregation patterns of markers in sets of recombinant inbred lines; this has been applied in several recent studies of the genetic basis of behavioral traits in mice (reviewed by CRABBE *et al.* 1994).

An alternative experimental strategy, the basis of the present study, is to follow frequency changes at marker loci in selection lines (DUMOUCHEL and ANDERSON 1968; GARNETT and FALCONER 1975; KEIGHTLEY and BULFIELD 1993; NUZHDIN *et al.* 1993). It is possible to estimate linked QTL effects from changes of frequencies of marker alleles after several generations of artificial selection on the quantitative trait of interest, and the task is made easier if the base population allele frequencies are known and the population is initially in a state of complete linkage disequilibrium (*e.g.*, derived from a cross between inbreds). The approach should reveal only tight linkage between markers and QTLs because loose associations will tend to become quickly broken down by recombination. In this paper, the genetic basis of the response to selection on body weight at 6 wk in mice is investigated using a set of replicated lines divergently selected for 21 generations whose base population was a cross between the inbred strains C57BL/6J and DBA/2J. At the end of this period, allele frequencies were measured at marker loci, the map locations of which were inferred elsewhere (DIETRICH *et al.* 1992, 1994). The presence of QTLs is indicated where regions of the genome show large differences in allele frequencies between high and low selection lines.

Corresponding author: Peter D. Keightley, Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JT, Scotland.
E-mail: p.keightley@edinburgh.ac.uk

[1] *Present address:* Institut für Grundlagen der Nutztierwissenschaften, Fachgebiet Züchtungsbiologie und molekulare Tierzüchtung, Humboldt Universität zu Berlin, Invalidenstrasse 42, Berlin, Germany.

[2] *Present address:* Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland.

A method based on maximum likelihood (ML) is used to estimate effects and map positions of QTLs relative to the markers based on genotype frequencies at one marker locus (KEIGHTLEY and BULFIELD 1993) and extended to allow interval mapping with several markers at a time.

MATERIALS AND METHODS

Mouse strains: The base population (generation 0) was an F_2 of the inbred strains C57BL/6J and DBA/2J obtained from the Jackson Laboratory, Maine in 1985. Selection was carried out on a within family basis on 6 wk weight for 21 generations. Six high and six low lines of eight pairs of parents each were maintained. A statistical analysis of the data from the selection experiment is described in detail elsewhere (HEATH *et al.* 1995).

Animals: Genomic DNA samples from a total of 93 animals were obtained from the low line replicates, and these constituted nearly a complete set of parents. Samples from 34 individuals from the high lines were available. In this case, the replicate assignments had to be inferred by a phylogenetic analysis based on marker genotypes (KEIGHTLEY 1994). The number of animals in the six high line replicates ranged from 3 to 9. The likelihood analysis takes account of the different numbers of individuals in the high and low line replicates. Although relatively few individuals are genotyped, the power to detect QTLs is comparable to a large experiment of a more conventional design (KEIGHTLEY and BULFIELD 1993).

Microsatellite marker locus typing: Individuals were typed at a total of 124 microsatellite marker loci (HEARNE *et al.* 1991; DIETRICH *et al.* 1992, 1994). PCR products were visualized under UV light with ethidium bromide staining after separation on 20 cm polyacrylamide gels. For each marker, the three genotypes could be distinguished unambiguously.

Likelihood analysis: A method based on Monte Carlo simulation was used to evaluate likelihood of observed genotype frequencies at one or more linked markers as a function of effect and map position of one or more QTLs. This extends to multiple loci a method to estimate QTL effects from frequency changes of markers under the assumption of complete linkage between a marker and a QTL (KEIGHTLEY and BULFIELD 1993). Haldane's mapping function (HALDANE 1919) was used to relate recombination probability to map distance. The procedure modeled directional selection at a QTL in order to generate the expected frequency distribution of linked marker genotypes. The within family selection practised in the experiment was approximated by modeling fertility selection in an idealized population. This was computationally more efficient and made the likelihood more straightforward to calculate. The accuracy of this approximation is investigated in RESULTS. To generate N progeny, N pairs of parents were sampled with replacement. The probability of sampling an individual was proportional to its relative fitness, which was 1 or $1 + s$ if homozygous or $1 + s/2$ if heterozygous. One offspring was generated per sample pair thus generating a Poisson distribution of family size. The selection coefficient s may be converted to a scale in units of the trait by applying the formula

$$a = 2s\sigma_P(1 - t)^{1/2}/i, \quad (1)$$

where a is the effect on the trait measured as the difference between the homozygotes, i is the intensity of selection, t is the correlation of phenotypic values of members of the same family and σ_P is the phenotypic standard deviation. This corrects a slight error in the formula given by KEIGHTLEY and BULFIELD (1993), which included a term n for family size, as this is implicit in the selection intensity term. In order to generate the appropriate amount of genetic drift of allele frequencies at QTLs and linked markers as expected for the selection lines, the population size was set to 23, the mean effective population size of the selection lines obtained from the pedigree by computing the mean relationship between individuals within lines at generation 20. For within family selection with equal representation among families, the effective population size is expected to be about 32, *i.e.*, twice the actual population size (FALCONER 1989), but was observed to be less than this due to the failure of some matings.

The simulation was started from an initial state with allele frequencies of 0.5, complete linkage disequilibrium and a QTL of effect s at some position relative to the markers and run for 20 generations with $N = 23$. The simulation was then run for one further generation, and a large number of progeny generated, from which a vector $\mathbf{p}$ of probabilities of the k marker genotype classes (where $k = 3^m$, and m is the number of markers) was computed. The likelihood of an observed vector $\mathbf{X}$ of numbers of individuals in the k genotype classes was computed for the simulation replicate j under the assumption of a multinomial distribution: $L_j(X_1 = x_1, \cdots X_k = x_k) = (n!/x_1! \cdots x_k!)p_1^{x_1} \cdots p_k^{x_k}$, where n is the number of individuals. The likelihood for each of the selection lines in the experiment was the average of L_j over a large number (typically 5×10^4) of simulation runs, and the overall log likelihood of the data was the sum of log likelihoods for individual independent replicates. Divergent selection of equal intensity in the high and low replicates was built in by assuming an opposite sign for s when computing likelihoods for high and low line replicates. The algorithm was computationally demanding at the time of writing, especially with data from more than one marker. The analysis was carried out for autosomal loci only as marker alleles on the X chromosome showed only small divergences in frequency between the high and low lines. In the analysis, map positions were assumed to be known with certainty, as there was no information available in the experiment to estimate these.

RESULTS

Marker allele frequency changes: Within family selection on body weight at 6 wk was carried out in six replicates in each direction. The high and low lines differed by about 9 g at generation 21 [four phenotypic standard deviations (SD)] or 45% of the mean of the inbreds, which hardly differ for body weight; and the realised heritability estimate was 22% (HEATH *et al.* 1995). Allele frequencies were measured at a total of 124 microsatellite marker loci (Table 1). Initially, markers were chosen for analysis from the ca. 4000 loci available (DIETRICH *et al.* 1994) such that there would be an even spacing of about 20 cM throughout the genome. Subsequently, whenever alleles at one or several markers in a region showed a frequency divergence (Δ_q) between the high and low lines of greater than about 0.4, further loci were typed in order to locate peaks in Δ_q. A genetic map with allele frequency changes along with log likelihood ratios for an analysis in which complete linkage between individual markers and a QTL was assumed is shown in Figure 1. The log likelihood ratio is the difference between the natural log maximum likelihood and the log likelihood for the QTL

TABLE 1

List of the microsatellite marker loci typed

Marker	Map position	Marker	Map position	Marker	Map position	Marker	Map position
D1Mit118	4.3	*D6Mit86*	0.0	*D12Mit46*	12.4	*D18Mit14*	10.2
D1Mit213	26.4	*D6Mit183*	13.5	*D12Mit2*	15.7	*D18Mit17*	14.8
D1Mit7	43.7	*D6Mit17*	20.2	*D12Mit158*	32.7	*D18Mit9*	27.1
D1Mit48	55.6	*D6Mit9*	28.0	*D12Mit7*	45.0	*D18Mit7*	36.2
D1Mit191	69.8	*D6Mit39*	40.5	*D12Mit79*	50.7		
D1Mit30	76.6	*D6Mit55*	42.9	*D12Nds2*	60.1	*D19Mit61*	11.0
D1Mit265	80.0	*D6Mit14*	61.1			*D19Mit13*	22.9
D1Mit57	86.7			*D13Mit3*	2.2	*D19Mit1*	43.2
D1Mit155	116.6	*D7Mit57*	6.8	*D13Mit91*	15.6		
		D7Mit55	12.3	*D13Mit9*	25.7	*DXMit89*	1.1
D2Mit32	13.4	*D7Mit193*	22.8	*D13Mit149*	39.2	*DXMit1*	29.9
D2Mit61	35.4	*D7Mit82*	25.1	*D13Mit77*	55.3	*DXMit117*	45.8
D2Mit14	48.9	*D7Mit84*	26.7	*D13Mit35*	61.3		
D2Mit17	56.7	*D7Mit62*	32.9				
D2Mit285	72.3	*D7Mit66*	45.4	*D14Mit1*	1.1		
D2Mit200	96.2	*D7Mit12*	65.3	*D14Nds1*	1.1		
				D14Mit50	5.0		
D3Mit54	4.6	*D8Mit63*	10.4	*D14Mit5*	24.6		
D3Mit21	14.7	*D8Mit31*	32.2	*D14Mit7*	46.5		
D3Mit51	27.1	*D8Mit200*	56.0	*D14Mit95*	58.0		
D3Mit189	38.3	*D8Mit56*	70.6				
D3Mit17	51.7			*D15Mit13*	0.0		
D3Mit19	68.6	*D9Mit67*	11.8	*D15Mit5*	18.3		
		D9Mit4	22.6	*D15Mit29*	34.0		
D4Mit101	4.1	*D9Mit146*	40.6	*D15Mit34*	53.9		
D4Mit55	17.3	*D9Mit20*	56.0	*D15Mit16*	61.7		
D4Mit139	23.0	*D9Mit18*	65.3				
D4Mit142	25.2			*D16Mit122*	6.8		
D4Mit81	26.3	*D10Mit3*	12.4	*D16Mit88*	10.2		
D4Mit185	29.7	*D10Mit42*	38.6	*D16Mit4*	27.1		
D4Mit76	46.3	*D10Mit11*	47.6	*D16Mit5*	34.9		
D4Mit16	49.0	*D10Mit180*	63.4	*D16Mit50*	41.7		
D4Mit54	60.3			*D16Mit70*	52.1		
D5Mit227	5.6	*D11Mit63*	2.2	*D17Mit46*	2.2		
D5Mit11	18.1	*D11JT4*	10.0	*D17Mit16*	6.7		
D5Mit18	34.0	*D11Mit20*	18.0	*D17Mit11*	10.0		
D5Mit24	47.5	*D11JT3*	30.0	*D17Mit49*	13.4		
D5Mit136	55.3	*D11Nds1*	39.0	*D17Mit66*	19.2		
D5Mit95	59.9	*D11Nds16*	47.0	*D17Mit3*	35.2		
D5Mit65	59.9	*D11JT147*	53.0	*D17Mit41*	47.6		
D5Mit98	68.8	*D11Nds8*	59.0				
D5Mit99	75.7	*D11Mit132*	62.2				
D5Mit222	78.2	*D11Mit48*	81.3				

Map positions (cM) are from 1994 release of the Research Genetics database (see DIETRICH *et al.* 1994), with the exception of markers on chromosome *11* with label JT which were obtained from HEARNE *et al.* (1991). Markers which were included after the initial mapping phase in order to saturate a region are underlined.

effect set to zero. Peaks of Δ_q ranging from about 10 cM (*e.g.*, chromosome *4*) to most of the chromosome (chromosome *9*) are present. The estimated QTL effects associated with the 11 regions where log likelihood ratios exceed 3, along with the identities of the markers involved, are shown in Table 2. Many loci of moderate effect apparently contributed to the selection response, but the additive genetic variance contributed by each locus was small. Four QTLs on chromosomes *4, 7, 9* and *11* previously detected by allele frequency changes at coat color and endogenous retrovirus markers (KEIGHTLEY and BULFIELD 1993) are also detected by linked microsatellite markers.

Expected distribution of effects under a null hypothesis: The presence of genetic linkage and an unknown distribution of QTL effects makes the setting of a likelihood ratio threshold at which to accept or reject a QTL-marker association a difficult problem in QTL mapping studies. Too stringent a threshold implies that many QTLs will be missed (type II statistical errors); and this

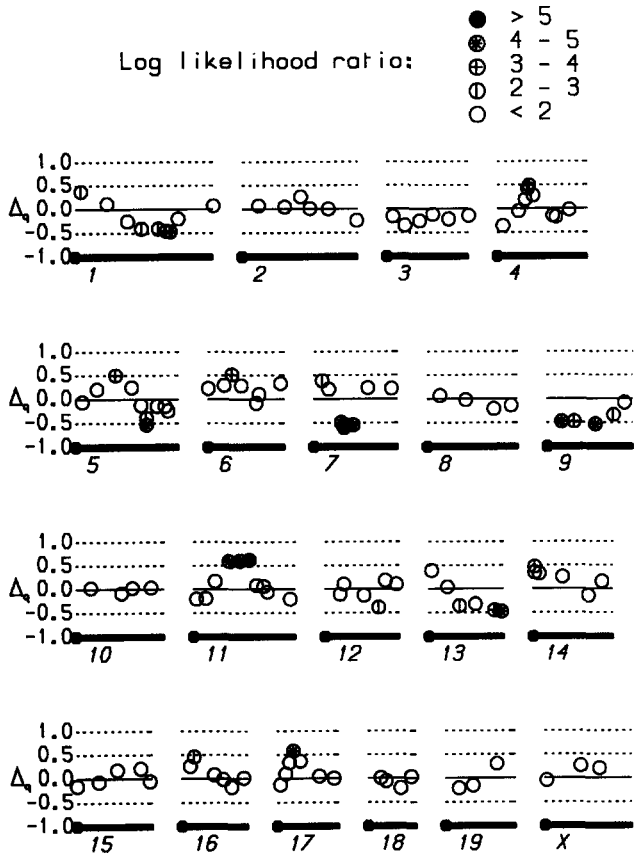

FIGURE 1.—Map showing the mean difference in allele frequency (Δ_q) averaged over replicates between high and low selection lines at microsatellite marker loci on 20 mouse chromosomes. A Δ_q value of 1.0, for example, would mean that the high lines were all fixed for the DBA/2J alleles and the low lines were fixed for the C57BL/6J alleles. The order of the markers along the chromosomes is the same as in Table 1. The natural log likelihood ratios for a model of complete linkage between a QTL and individual markers are indicated.

may be as undesirable as setting too low a threshold, which will tend to generate false positives (type I errors). To attempt to address this problem in the context of the present study, Monte Carlo simulation of the drift in frequency of alleles at a neutral marker was used to generate data sets with the same structure as the experiment. These data sets were analyzed to infer the distribution of effects associated with unlinked markers with zero effect on the trait (Figure 2). Also shown is the observed distribution for markers at least 10 cM apart and includes markers corresponding to peaks in Δ_q. There is clearly a large excess of markers associated with effects in the tails of the distribution. The distribution under the null hypothesis suggests strong support for the presence of QTLs with effects greater than 0.2 phenotypic SD in six regions (chromosomes *5, 7, 9, 11, 13* and *17*), whereas the expected number of effects exceeding this magnitude is about 0.3. There is rather weaker support for the presence of QTLs on chromosomes *4, 5, 6* and *14*. The presence of QTLs on chromosomes *4* and *9* had previously been

TABLE 2

QTL effects identified

Marker	Log likelihood ratio	Effect (SD)	Heritability (%)
D1Mit30	3.3	−0.17	0.3
D4Mit81	3.3	0.18	0.4
D5Mit18	3.2	0.18	0.4
D5Mit65	4.3	−0.21	0.5
D6Mit17	4.0	0.19	0.4
D7Mit82	6.9	−0.28	0.9
D9Mit146	4.4	−0.21	0.5
D11Nds16	5.7	0.24	0.6
D13Mit35	4.1	−0.20	0.4
D14Mit1	3.0	0.17	0.3
D17Mit49	4.9	0.22	0.7

Natural log likelihood ratios, effects in phenotypic standard deviation units and heritabilities associated with microsatellite marker loci which showed the highest divergences of allele frequency. Estimates of heritabilities associated with markers included a term for the sampling variances of the estimates obtained from the rate of change of likelihood near the maximum (KEIGHTLEY and BULFIELD 1993).

inferred from consistent changes of allele frequencies at the *brown* and *dilute* coat color loci over the course of the selection experiment (KEIGHTLEY and BULFIELD 1993).

In order to investigate the effect of background variation for the trait, background variation in fitness was modeled with 32 unlinked genes of equal effect such that the heritability for fitness would be similar to that induced by artificial selection in the experiment. In this case, the population size was set to 32, a value which accounts for the expected reduction in effective popula-

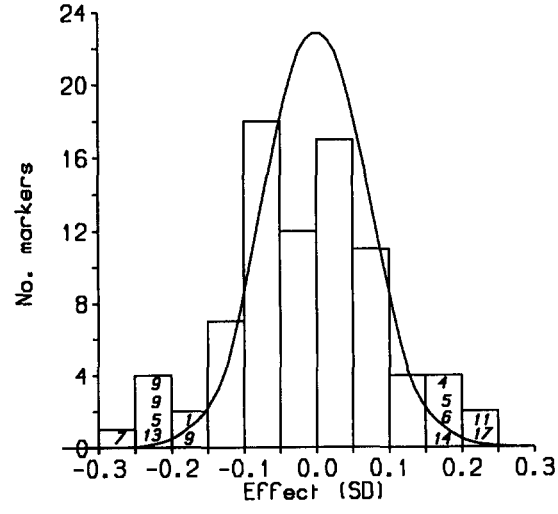

FIGURE 2.—Curve showing the expected distribution of ML estimates of effects (in phenotypic SD units) associated with neutral markers, generated by Monte Carlo simulation. The histogram shows the observed distribution of ML estimates of effects associated with a subset of quasi-independent markers at least 10 cM apart. The figures inside the histogram refer to the chromosome in which the putative QTLs are located.

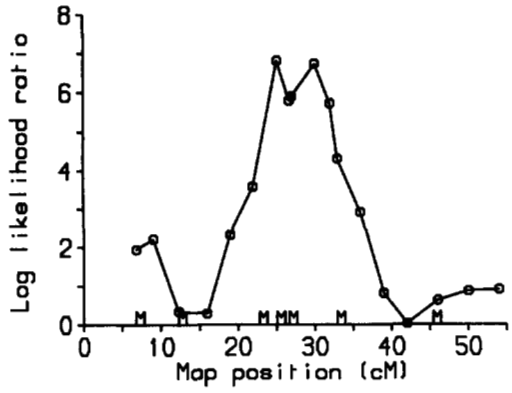

FIGURE 3.—Natural log likelihood ratio for data of chromosome 7 evaluated at various positions for one QTL flanked by pairs of marker loci based on interval mapping. The map locations of the markers are indicated by M.

tion size as a consequence of genetic variance in fitness. This figure was derived by applying a formula of SANTIAGO and CABALLERO (1995). Likelihoods computed with the inclusion of background variation for fitness were found to be essentially the same as for the absence of background variation but with the smaller population size.

Interval mapping: The presence of a number of sharp peaks in the change of marker allele frequency (Figure 1) suggests that single genes of moderate effect on the trait may be present. A variant of the interval mapping approach (LANDER and BOTSTEIN 1989) was used to generate support limits for the location of QTLs near these peaks. For example, with data from chromosome 7, likelihood ratios for the presence of one QTL flanked by pairs of markers were computed at different points along the chromosome for pairs of markers (Figure 3). If a drop in natural log likelihood of 2 from the ML is taken as the support limit (asymptotically equivalent to a 95% confidence limit), the support limits for the position of the QTLs are 23 and 33 cM relative to marker *D7Mit82* (map position 25.1 cM). Support intervals for several QTLs are in the region of 10 cM, but some (*e.g.*, chromosome *9*) include most of the chromosome (Table 3). Note that the ML estimates of QTL map positions tend to be close to the positions of markers showing the greatest change of allele frequency.

Because of the high computing requirements, only limited analyses could be carried out in which three markers were fitted simultaneously. For example, likelihood surfaces as a function of the effect and map position of one QTL for regions of chromosomes *9* and *17* covered by three markers are shown in Figure 4. In the case of chromosome *17*, there is a peak in the likelihood surface close to the marker *D17Mit49* (which shows the highest change in allele frequency), but the support interval for QTL position is not much different from that computed by interval mapping (Table 3). In the case of chromosome *9*, there is a long saddle in the likelihood surface covering a large part of the chromosome, with peaks in likelihood between the markers and dips at the markers themselves.

More than one QTL in a region: The computing requirements for comparison of likelihoods of one or two QTLs in a region were still higher than the above, but limited analysis was nonetheless possible. The downhill simplex algorithm (NELDER and MEAD 1965) was used to maximize likelihood as a function of two (rather than one) QTL effects and map positions with data of the three chromosome *17* markers shown in Figure 4A. The likelihood appeared to converge to a similar value to the ML value for a single QTL, and the positions of each QTL converged to the assumed map position of the marker *D17Mit49*, where the highest change in marker allele frequency occurs. For chromosome *9*, where there is a significant divergence of allele frequency over much of the chromosome, likelihood was computed for a model with two QTLs and markers as shown in Figure 4B. As in the analysis of chromosome *17* data, a difference between the likelihoods of the one and two QTLs models was not detected. The highest likelihood occurs with QTLs midway between markers *D9Mit67* and *D9Mit146*, where there is a plateau in the change of marker allele frequency. For chromosome *5*, where there is a switch in the direction of change of marker allele frequency across the chromosome, a model with two QTLs of opposite sign located at the markers showing peaks of Δ_q gave a significantly better fit than single QTL at either marker (log likelihood ratios were 3.5 and 4.4).

Candidate genes: A class of loci intimately involved in growth control is connected with the somatotropic axis (FROESCH *et al.* 1985). There are also loci known primarily for their major effects on growth or level of obesity (LYON and SEARLE 1989). Map positions of some of these candidate genes were inferred relative to the microsatellite map from data of COPELAND *et al.* (1993), who mapped a number of microsatellite loci on a gene-based map, derived from the typing an interspecies backcross (COPELAND and JENKINS 1991). Information on map positions of the candidate genes relative to the markers was also obtained from the reports of single chromosome committees (The Encyclopaedia of the Mouse Genome, Release 3). Although the absolute map locations of the microsatellite markers provided by DIETRICH *et al.* (1992, 1994) do not in general agree with the above sources, the relative positions are in reasonable agreement. Log likelihood ratios for the presence of a QTL flanked by microsatellite markers flanking candidate genes of the somatotropic axis and other candidate loci are shown in Table 4. *Igf1r* (the insulin-like growth factor 1 receptor locus), on chromosome *7*, maps to a region strongly associated with body size in the selection lines, as does another linked locus affecting size, *tub*. On chromosome *6* an interval containing the *lit* locus also shows a significant association.

TABLE 3

Support intervals for QTL map positions

Chromosome	Reference marker	Map position	ML QTL position	Support limits		Log likelihood ratio
				Lower	Upper	
1	*D1Mit30*	76.6	76	46	84	3.4
4	*D4Mit81*	26.3	26	24	30	3.3
5	*D5Mit65*	59.9	60	57	64	4.2
5	*D5Mit18*	34.0	35	22	45	3.4
6	*D6Mit17*	20.2	22	15	26	4.3
7	*D7Mit82*	25.1	25	23	33	6.8
9	*D9Mit146*	40.6	32	12	50	6.3
11	*D11Nds16*	47.0	45	29	49	6.2
13	*D13Mit35*	61.3	59	29	Telomere	4.7
14	*D14Mit1*	1.1	0	0	22	2.8
17	*D17Mit49*	13.4	14	11	18	4.8

Maximum likelihood estimates of map positions (cM) of QTLs relative to microsatellite markers, along with support intervals.

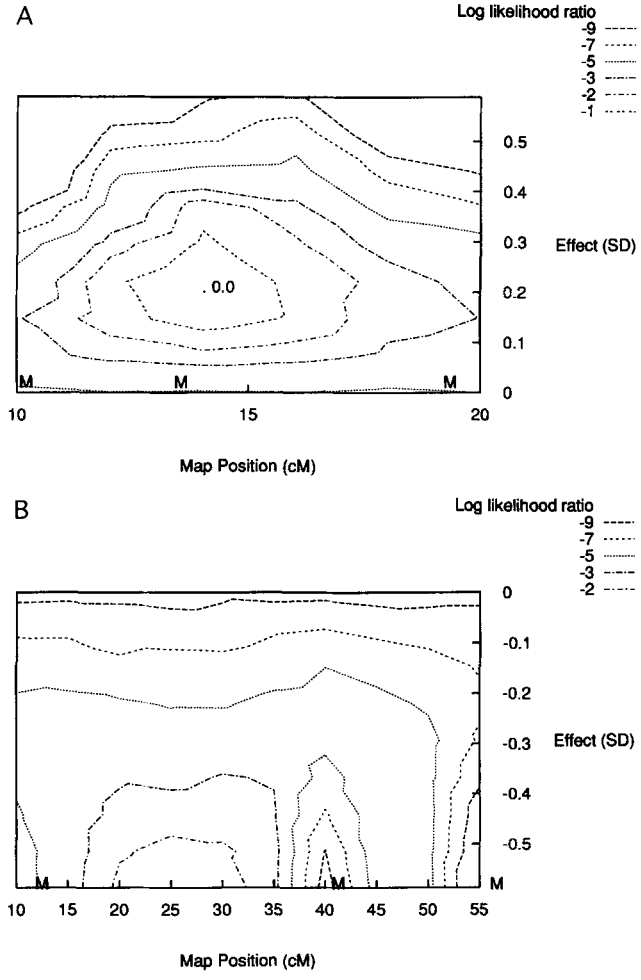

FIGURE 4.—Surfaces showing the likelihood ratio for data with three markers fitted simultaneously as a function of map position (cM) and QTL effect (in phenotypic SD units). Map positions of the markers are indicated by M. (A) Chromosome *17*. Markers fitted: *D17Mit11, D17Mit49, D17Mit66*. (B) Chromosome *9*. Markers fitted: *D9Mit67, D9Mit146, D9Mit20*.

Other candidate loci do not show significant associations. Note that the likelihood ratio for the region fitted on chromosome 7 is higher than obtained for the single marker analysis (Table 2) or from interval mapping (Figure 3), presumably because, in this case, data from relatively widely spaced markers *both* showing large changes of frequency were analyzed. A similar effect occurs in the analysis of chromosome *9* data with three markers at a time (Fig. 4b).

Performance of fertility selection approximation: The maximum likelihood estimation of QTL effects was carried out under a fertility selection model, whereas within family selection was practised in the experiment. The accuracy of this approximation was tested by simulating data sets using within family selection or fertility selection with equivalent effective population sizes and selection coefficients [assuming (1)], then analyzing the data under the fertility selection model. Mean estimates of selection coefficients and log likelihood ratios for a range of simulated selection coefficient values as a function of the overall change of marker allele frequency are shown in Figure 5. There is a slight tendency for the within family selection data sets to lead to higher estimates of selection coefficients and log likelihood ratios. The reason for this appears to be a higher level of heterozygosity for within family selection with equal contributions, leading to a higher change of allele frequency even in the initial generations. Likelihood ratios are very similar for a given change of allele frequency.

DISCUSSION

The data suggest that a large number of genetic differences between two inbred strains contributed to a substantial response to artificial selection on body weight. This genetic variation was hidden in the base population, as the inbreds hardly differ for the trait. There is strong support for the presence of seven QTLs

TABLE 4

Candidate genes

Chromosome	Candidate gene	Flanking markers	Log likelihood ratio
4	*db*	*D4Mit76, D4Mit16*	0.6
6	*ob*	*D6Mit86, D6Mit183*	1.6
6	*lit*	*D6Mit183, D6Mit17*	4.2
7	*Igf1r*	*D7Mit82, D7Mit62*	9.6
7	*tub*	*D7Mit62, D7Mit66*	5.2
7	*Igf2*	*D7Mit12*[a]	0.7
7	*ad*	*D7Mit12*[a]	0.7
10	*hg*	*D10Mit42, D10Mit11*	0.0
10	*Igf1*	*D10Mit42, D10Mit11*	0.0
11	*Gh*	*D11Mit132, D11Mit48*	0.8
15	*Ghr*	*D15Mit13*[b]	0.6
16	*dw*	*D16Mit5, D16Mit50*	0.2
17	*Igf2r*	*D17Mit46*[c]	0.3

Candidate genes, flanking markers, and natural log likelihood ratios for the presence of a QTL in the interval between the markers.

[a] Interval from marker to telomere; one marker fitted only.

[b] Marker and candidate gene at same position or very close: one marker fitted only.

[c] Interval from centromere to marker: one marker fitted only.

with moderate effects and rather weaker support for the presence of an additional four. In only one case is there strong evidence of more than one QTL on a chromosome. Surprisingly, the QTLs identified account for only about one-quarter of the genetic variation in the base population. If genetic variance estimates associated with these and all the other markers at least 10 cM apart are summed (which presumably accounts for some of the variance associated with lesser QTLs), only about one-third of the genetic variance in the base population is accounted for. There are several possible explanations for this underestimation. It is possible that some peaks of change in allele frequency have been missed; but a serious effort has been made to cover the genome, and simulations suggest that QTLs of large effect are unlikely to have been missed (KEIGHTLEY and BULFIELD 1993). A more likely explanation is that the additive model under which effects are estimated is inadequate. If genes act in a nonadditive manner, which can be inferred due to the presence of heterosis in the F_1 and changes of variance that do not fit with an infinitesimal model (HEATH *et al.* 1995), the additive model could underestimate the variances associated with the markers. An attempt was made to estimate dominance effects of markers, but likelihood as a function of degree of dominance was found to be nearly flat. There is also the possibility of coupling between QTL with effects in the same direction, which would give the appearance of large amounts of genetic variance in the base population, but changes of allele frequencies at marker loci would be less than expected from the selection response. This appears to be contrary to observation of an increase in genetic variance as a consequence of directional selection in these lines (HEATH *et al.* 1995). Some regions of the genome covering large parts of individual chromosomes (*e.g.*, chromosomes *9* and *11*) showed changes in frequency in the same direction that were greater than would be expected under a drift model, but there is inadequate power to distinguish the presence of one or several QTL. A finer scale mapping strategy is required to answer this question.

The analysis assumes that the base population marker allele frequencies are known (*i.e.*, 0.5 in the F_1), and frequency changes occur subsequently only due to drift or by hitchhiking to QTL alleles under selection. The phenomenon of segregation distortion can occur in mice (SILVER 1993) and could also lead to directional changes of marker allele frequencies; but its presence would not seriously affect the results of the analysis as most of the information to estimate QTL effects comes from differences in frequencies between high and low selection line replicates.

As many as one-quarter of catalogued mouse mutants have a pleiotropic or main effect on some aspect of growth (KEIGHTLEY and HILL 1992). In contrast to bristle number in *Drosophila*, for which the number of candidate genes is more limited (LONG *et al.* 1995), numbers of candidate genes for growth in mammals is large. Tests carried out to determine if any of the QTLs map close to candidate genes of the somatotropic axis (*Igf1, Igf1r, Igf2, Igf2r, Gh, Ghr*) or map close to other candidate genes for which mutants of large effect are known in mice suggest that one QTL on chromosome 7, which changes body weight by about 3%, maps close to *Igf1r* and *tub*, while a QTL of rather smaller effect maps close to *lit*. The regions in question cover about 20 cM, so mapping to a much finer scale is necessary to determine if a variant at the candidate is causal. The need for fine scale mapping is illustrated by a recent case of mapping

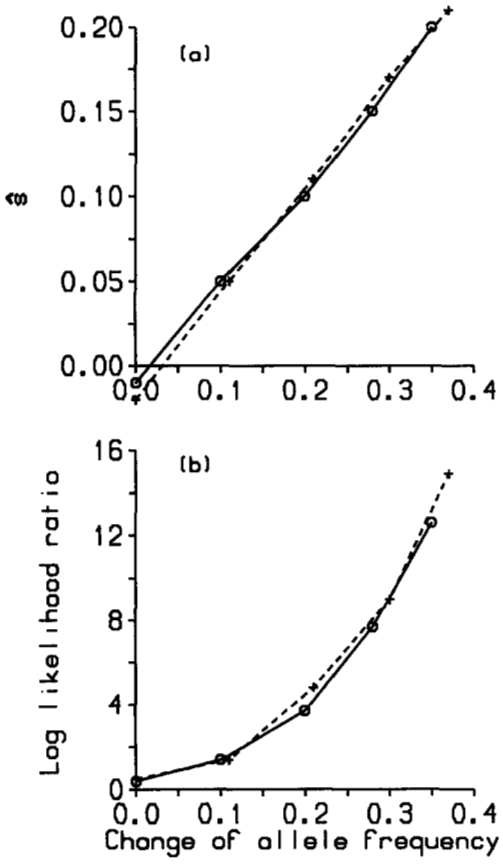

FIGURE 5.—Performance of the fertility selection approximation. Replicated data sets of genotype frequencies for one marker completely linked to a QTL were generated under models of fertility selection or within family selection. Gene effects for the trait under within family selection as a function of s were obtained assuming equation (1). The mean selection coefficient associated with the marker $\hat{s}$ and log likelihood ratios were estimated with the fertility selection model and are plotted against the observed change of allele frequency to generation 20. The initial gene frequency was 0.5, and selection coefficient values were 0, 0.05, 0.1, 0.15 or 0.2. The fertility selection model used 32 parents, and the within family selection model used eight pairs of parents ($N = 16$; $N_e \approx 32$) and 48 progeny per generation ($i = 0.848$). For each simulation data set, 12 sublines were generated. +---+, within family selection; O——O, fertility selection.

of a mutation with a large positive effect on growth in mice, the high growth gene (*hg*) (BRADFORD and FAMULA 1984). This mutation originally mapped to a region on chromosome *10* close to the *Igf1* locus, but subsequent fine scale mapping showed that the *hg* mutation is not an allele at the *Igf1* locus (HORVAT and MEDRANO 1995). The candidate gene approach, however, shows promise in identifying variants responsible for genetic variation in bristle score of Drosophila in natural populations (MACKAY and LANGLEY 1990; LAI *et al.* 1995; LONG *et al.* 1995).

It can be concluded that the QTLs identified in the present study are mostly alleles of genes whose functions are as yet unknown. Furthermore, different QTLs tend to be involved in different populations. Associations have been found between growth in mice and the *Gh* and *Igf2* loci (WINKELMAN and HODGETTS 1992), neither of which show associations in the present study. In a preliminary report of a mapping study for QTLs involved in 10 wk body weight and fatness in mice, a QTL of major effect on chromosome *2* was found (POMP *et al.* 1994), where there is clearly no major effect in the present study. A further QTL on chromosome *11* (POMP *et al.* 1994) may map close to that detected in the present population.

If selection lines are available, mapping of QTLs by following frequencies of marker alleles is a powerful approach, as QTLs with relatively small effects can be detected by genotyping small numbers of individuals. The availability of microsatellite markers that can be easily typed for large sample sizes makes this less of an advantage, however. The method does not necessarily depend on knowing the base population gene frequencies, because in principle, these can be estimated simultaneously in the model. Most of the information to detect QTLs comes from the divergence of allele frequency between selection lines. The mapping is on a relatively fine scale because there are many more opportunities for recombination between markers and QTLs compared to conventional designs, but this is not necessarily an advantage for an initial scan of the genome. The ultimate molecular identification of the genes involved will require mapping strategies that bridge the gap between physical maps of the mouse genome currently under construction and the 10–20 cM resolution achieved by the present method.

We are grateful to S. C. HEATH and A. L. ARCHIBALD for helpful advice and to W. G. HILL, T. F. C. MACKAY and two anonymous referees for constructive criticisms. We thank the Biotechnology and Biological Sciences Research Council, the Royal Society (PDK) and the Leopoldina (TH) for support.

LITERATURE CITED

BRADFORD, G. E., and T. R. FAMULA, 1984 Evidence for a major gene for rapid post-weaning growth in mice. Genet. Res. **44:** 293–298.

COPELAND, N. G., and N. A. JENKINS, 1991 Development and applications of a molecular genetic linkage map of the mouse genome. Trends Genet. **7:** 113–118.

COPELAND, N. G., N. A. JENKINS, D. J. GILBERT, J. T. EPPIG, L. J. MALTAIS *et al.*, 1993 A genetic linkage map of the mouse: current applications and future prospects. Science **262:** 57–66.

CRABBE, J. C., J. K. BELKNAP and K. J. BUCK, 1994 Genetic animal models of alcohol and drug abuse. Science **264:** 1715–1723.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper and Row, New York.

DIETRICH, W., H. KATZ, S. E. LINCOLN, H.-S. SHIN, J. FRIEDMAN *et al.*, 1992 A genetic map of the mouse suitable for typing interspecific crosses. Genetics **131:** 423–447.

DIETRICH, W., J. C. MILLER, R. G. STEEN, M. MERCHANT, D. DAMRON *et al.*, 1994 A genetic map of the mouse with 4,006 simple sequence length polymorphisms. Nature Genet. **7:** 220–245.

DUMOUCHEL, W. H., and W. W. ANDERSON, 1968 The analysis of selection in experimental populations. Genetics **58:** 435–449.

FALCONER, D. S., 1989 *Introduction to Quantitative Genetics,* Ed. 3, Longman, London.

FRIEDMAN, J. M., and R. L. LEIBEL, 1992 Tackling a weighty problem. Cell **69:** 217–220.

FROESCH, E. R., C. SCHMID, J. SCHWANDER, and J. ZAPF, 1985 Action of insulin-like growth factors. Annu. Rev. Physiol. **47:** 443–467.

GARNETT, I., and D. S. FALCONER, 1975 Protein variation in strains of mice differing in body size. Genet. Res. **25:** 45–57.

HALDANE, J. B. S., 1919 The combination of linkage values and the calculation of distance between loci of linked factors. J. Genet. **8:** 299–309.

HALEY, C. S., and S. A. KNOTT, 1992 A simple method for mapping quantitative trait loci in line crosses using flanking markers. Heredity **69:** 315–324.

HEARNE, C. M., M. A. MCALEER, J. M. LOVE, T. J. AITMAN, R. J. CORNALL *et al.,* 1991 Additional microsatellite markers for mouse genome mapping. Mamm. Genome **1:** 273–282.

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.

HORVAT, S., and J. F. MEDRANO, 1995 Interval mapping of high growth (*hg*), a major locus that increases weight gain in mice. Genetics **139:** 1737–1748.

JANSEN, R. C., and P. STAM, 1994 High resolution of quantitative traits into multiple loci via interval mapping. Genetics **136:** 1447–1455.

KEIGHTLEY, P. D., 1994 Recovery of phylogenetic information with microsatellite markers. Mouse Genome **92:** 683–685.

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.

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

LAI, C., R. F. LYMAN, A. D. LONG, C. H. LANGLEY and T. F. C. MACKAY, 1995 Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster.* Science **266:** 1697–1702.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LONG, A. D., S. L. MULLANEY, L. A. REID, J. D. FRY, C. H. LANGLEY *et al.,* 1995 High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster.* Genetics **139:** 1273–1291.

LYON, M., and A. SEARLE, 1989 *Genetic Variants and Strains of the Laboratory Mouse,* Ed. 2, Oxford Univ. Press.

MACKAY, T. F. C., and C. H. LANGLEY, 1990 Molecular and phenotypic variation in the achaete-scute region of *Drosophila melanogaster.* Nature **348:** 64–66.

NELDER, J. A., and R. MEAD, 1965 A simplex method for function minimization. Computer J. **7:** 308–313.

NUZHDIN, S. V., P. D. KEIGHTLEY, and E. PASYUKOVA, 1993 The use of retrotransposons as markers for mapping genes responsible for fitness differences between related *Drosophila melanogaster* strains. Genet. Res. **62:** 125–131.

POMP, D., M. A. CUSHMAN, S. C. FOSTER, D. K. DRUDIK, M. FORTMAN *et al.,* 1994 Identification of quantitative trait loci for body weight and body fat in mice. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production,* **21:** 209–212. Univ. Guelph., Guelph, Ontario, Canada.

SANTIAGO, E., and A. CABALLERO, 1995 Effective size of populations under selection. Genetics **139:** 1013–1030.

SAX, K., 1923 The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris.* Genetics **8:** 552–560.

SILVER, L. M., 1993 The peculiar journey of a selfish chromosome: mouse *t* haplotypes and meiotic drive. Trends Genet. **9:** 250–254.

THODAY, J. M., 1961 Location of polygenes. Nature **191:** 368–370.

WINKELMAN, D. C., and R. B. HODGETTS, 1992 RFLPs for somatotropic genes identify quantitative trait loci for growth in mice. Genetics **131:** 929–937.

ZENG, Z.-B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Communicating editor: Z-B. ZENG