Genet. Res., Camb. (1997), **70**, *pp.* 125–133 *With* 4 *figures.* Printed in the United Kingdom © 1997 Cambridge University Press

Mapping quantitative trait loci for body weight on the X chromosome in mice. II. Analysis of congenic backcrosses

KELLIE A. RANCE[1], SIMON C. HEATH[2] AND PETER D. KEIGHTLEY[1]*
[1] *Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland, UK*
[2] *Department of Statistics, University of Washington, Seattle, Washington, USA*

(*Received 14 February 1997 and in revised form 27 May 1997*)

Summary

In a QTL mapping study with an F_2 population of mice, we have shown that one or more sex-linked factors account for a large part of the divergence between mouse lines selected for high and low body weight. Here, we describe a study undertaken to map the putative X-linked quantitative trait loci (QTLs) by backcrossing segments of chromosome from the high line onto an inbred line derived from the low line, thereby removing possible contributions from the autosomes and linked segments of the X chromosome. Sublines containing a regional at the proximal end of the X chromosome were found to be associated with large differences in body weight, and to account for almost all the difference between the lines. A Markov chain Monte Carlo based multipoint linkage analysis incorporating the available marker and phenotypic information from the backcross pedigree was used to map the QTL to a region of about 6 cM. There was no evidence for QTLs elsewhere on the chromosome. The estimated QTL effect is approximately 20% of mean body weight in males and females at 10 weeks. From results obtained from this study and the accompanying F_2 analysis, we conclude the presence of a single factor for body weight localizing to about position ($\pm$SE) $26{\cdot}4 \pm 1{\cdot}2$ cM on the X chromosome, which increases body weight by approximately 18% at 10 weeks. A strategy to positionally clone the QTL is discussed.

1. Introduction

Previous analysis of the genetic differences between mouse lines selected for growth rate has shown evidence for a large X chromosome effect accounting for approximately 25% of the divergence between the lines (Hastings & Veerkamp, 1993; Veerkamp *et al.*, 1993; Rance *et al.*, 1994). An accompanying study (Rance *et al.*, 1997) has indicated the presence of a major locus for body weight at approximately 23 cM (95% confidence interval) on the X chromosome, with an estimated effect of 5·2 g (about 20%) on 10 week weight in both males and females.

Simulation studies have shown that F_2 and first backcross (BC1) experimental designs, often employed in QTL mapping experiments, do not allow the mapping of quantitative trait loci (QTLs) with great precision (van Ooijen, 1992; Darvasi *et al.*, 1993), largely due to genetic noise from linked and unlinked QTLs elsewhere on the genome, and due to the lack of recombination events. Analytical methods for mapping QTLs have been developed to remove genetic noise from linked and unlinked QTLs by fitting linked or unlinked markers as cofactors in the analysis (Jansen, 1993, 1996; Zeng, 1994). These methods require segregating populations to be genotyped at markers spanning the entire genome in order to allow the removal of background genetic noise. An experimental alternative to this analytical procedure, used in the present study, is to progressively backcross regions of the genome of interest onto an inbred background.

A progressive backcross approach has been employed to study the genetic basis of variation in bristle number in *Drosophila* (Shrimpton & Robertson, 1988*a*, *b*), involving the use of morphological markers to maintain specific regions of the third chromosome on a common genetic background, and to generate several congenic lines. This experimental design can be more powerful than a simple F_2 or BC1 population, as the congenic lines can be assumed to be genetically identical in all regions of the genome unlinked to the congenic region of interest. The objective of the experiment described in this paper was to backcross

* Corresponding author. Tel: +44 (0)131 650 5443. Fax: +44 (0)131 667 3210. e-mail: p.keightley@edinburgh.ac.uk.

specific regions of the high line X chromosome onto an inbred low line background, as a means of localizing the X-linked QTL, and to provide a basis for fine-scale genetic mapping.

2. Materials and methods

(i) *PCR amplification of microsatellites*

Genomic DNA was extracted from spleen or ear clip samples by phenol:chloroform extraction. The polymerase chain reaction (PCR) amplification protocol used to genotype animals from BC generations 1–3 was based on the method described by Dietrich *et al.* (1992). Genotyping of animals in BC generations 4 and 5 was carried out with improvements described by Routman & Cheverud (1994), in which the 55 °C annealing temperature was reduced to 50 °C to improve the signal/noise ratio, and with the addition of 1/5 volume of loading dye (60% sucrose, 1·0 mM cresol red) prior to PCR amplification.

(ii) *Backcross population*

The mouse lines used in this study (the P6 lines) were divergently selected for body weight for more than 50 generations. At generation 50 of selection, the mean body weights were 49·3 g and 17·3 g in the high and low lines respectively (Mbaga, 1996). The origins of the lines have been described previously by Sharp *et al.* (1984) and Hastings & Hill (1989). To initiate the backcross lines, females from generation 51 of selection of the P6 high line and males from the P6 low inbred line, at generation 6 of full sib inbreeding (set up at generation 45 of selection), were mated to produce an F_1 population. F_1 females were mated to inbred low line males (generation 7 of inbreeding) to form generation 1 of the progressive backcross population (BC1):

P6 high line (gen. 45) × P6 low inbred (P6 LI) (gen. 6)
↓
F1 × P6 LI (gen. (7)
↓
BC1 × P6 LI (gen. 8)
↓
BC2 × P6 LI (gen. 9)
↓
BC3 × P6 LI (gen. 10)
↓
BC4 × P6 LI (gen. 11)
↓
BC5 × P6 LI (gen. 12)

Following DNA extraction from all the parents and the BC1 females, these individuals were genotyped at the loci shown in Table 1 (with the exception of *DXMit187*), which span the length of the X chromosome. The genotypes were used to find marker alleles in the BC1 females which could be traced unequivocally to the high selection line or the inbred low line. Female BC1 individuals were selected to be dams of the BC generation 2, where the markers genotyped were fully informative.

Table 1. *Microsatellite markers used to control the progressive backcrossing*

Marker	Map position (cM)[a]
DXMit55	0·0
DXMit187	12·2
DXMit50	19·8
DXMit25	33·7
DXMit16	41·9
DXMit79	52·6
DXMit38	56·0
DXMit121	71·4
DXMit31	72·0

[a] Map positions are from the accompanying F_2 study (Rance *et al.*, 1997).

The objective of the mating scheme was to maintain segments of the high line X chromosome in a heterozygous state in the BC population dams (using markers spanning the length of the X chromosome). To do this, all backcross individuals were typed at the loci listed in Table 1. An example of the conservation of an X chromosome segment is shown in Fig. 1. Assuming no double crossovers between flanking markers, the chromosome segment between markers B and D would be inherited from the high selection line, and marker alleles a and e inherited from the inbred low line. The proportion of the high line X chromosome between the markers a and B, and D and e, would also tend to decline with each generation of backcrossing.

The mating scheme was designed to establish a number of lines carrying overlapping segments of the X chromosome inherited from the high selection line, carried on the genetic background of the inbred low line. It should be noted that uniform segments were not obtained until later generations. In BC generations 3, a fully informative marker *DXMit*187 at 12·2 cM (Table 1) was added to the panel of eight markers used previously. The data collected during the population maintenance are summarized in Table 2.

(iii) *Effects associated with chromosome segments*

An analysis was carried out to estimate effects associated with carrying segments of the high line X chromosome flanked by known marker loci. The numbers of individuals carrying each segment were very small within each generation, because uniform segments were not established until the final generation

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

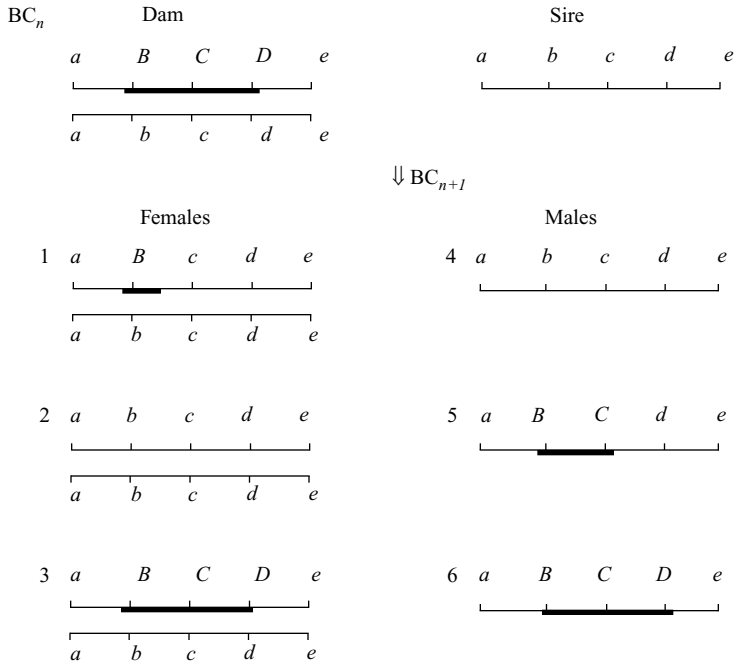

Fig. 1. An example of the maintenance of a high line X chromosome segment on the inbred low line background, with some examples of the possible backcross progeny.

Table 2. *Numbers of successful matings and individuals weaned in each generation of the backcross, and the traits recorded*

		Males			Females		
Generation	No. of litters	No. weaned	6 week weight	10 week weight	No. weaned	6 week weight	10 week weight
F_1	5	19[a]	✓		23	✓	
BC1	15	62[a]	✓	✓	84	✓	✓
BC2	11	—			57	✓	
BC3	14	90	✓	✓	87	✓	✓
BC4	17	96	✓	✓	83	✓	✓
BC5	22	123	✓	✓	112	✓	✓

[a] DNA available, but no genotypes recorded.

Table 3. *Means and variances of 6 and 10 week body weights for the F_1 and BC generations (corrected for the fixed effects and covariates)*

	6 week weight (g)		10 week weight (g)	
Generation	Mean	Variance	Mean	Variance
F_1	25·0	5·3	—	—
BC1	18·4	5·3	22·1	6·7
BC2	15·1	1·3	—	—
BC3	18·2	4·0	22·7	5·7
BC4	17·6	3·1	21·5	4·9
BC5	17·3	3·1	21·3	3·4

of backcrossing, so information from generations 3, 4 and 5 was combined. To attempt to account for changes in body weight observed over the generations analysed (see Table 3), a fixed effect of generation number was fitted. To estimate segment effects, the data were analysed in Genstat 5.3 (Genstat 5 Committee, 1993) using the following model:

$$Y_{ijklmn} = \mu + S_i + D_j + \beta \cdot W_{jk} + SEG_n + GN_l + e_{ijklmn},$$

where Y_{ijklmn} are observations of 6 or 10 week weight on the mth individual, μ is the overall mean, S_i is the fixed effect of the ith sex, D_j is the random common environmental effect associated with jth dam, β is the linear regression coefficient of Y on the number weaned (W_{jk}) from the kth parity litter of the jth dam, SEG_n is the fixed effect of the nth segment, GN_l is the fixed effect of the lth generation of origin, and e_{ijklmn} is the residual error associated with the mth individual.

(iv) *Analysis of the backcross population using all pedigree information*

An objective of the backcrossing strategy was to obtain an estimate for the X-linked QTL position and effect, which will provide information independent of

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

that already obtained from the analysis of the F_2 population (Rance *et al.*, 1997). To enable all the information collected in the different generations to be included in the analysis, the pedigree structure of the population needs to be accounted for. In the analysis of an F_2 or a simple first generation backcross population, using maximum likelihood or multiple regression methods (Lander & Botstein, 1989; Haley & Knott, 1992; Martinez & Curnow, 1992), the parental genotypes are assumed to be known, i.e. the founder lines are assumed to be fixed at different QTL alleles, and all the F_1 parents are heterozygous at the QTL. As a result of this assumption, the F_2 or first generation backcross individuals can be considered independently. The pedigree structure of the progressive BC population implies that all animals cannot be considered to be independent (with the exception of animals within the first generation of backcrossing). The individuals within generations are dependent on common ancestors in previous generations, resulting in dependencies across generations, and individuals from different litters may have ancestors in common in the previous generation, leading to dependencies within generations.

Another problem with dealing with more complicated pedigrees than result from F_2 or first generation backcross designs is that the markers are not equally informative. The markers flanking an interval do not provide all the information about the likelihood of a QTL within that interval, so to extract all the information about the position of a QTL on the chromosome a multipoint analysis simultaneously fitting all markers should be performed.

(v) *Use of MCMC for genetic analysis*

Accounting for missing genotype information in a genetic analysis is done by summation over all possible genotype configurations consistent with the data. There are two known exact algorithms for this summation: the peeling algorithm, proposed by Elston & Stewart (1971) and generalized by Cannings *et al.* (1978), and the Lander–Green algorithm (Lander & Green, 1987). Both methods are restricted in the types of problems they can be used for. The complexity of peeling is very roughly linear in pedigree size, but exponential in the number of loci being considered, and is thus limited to a relatively small number of loci. The reverse is true for the Lander–Green algorithm, which is roughly linear in the number of loci, but exponential in pedigree size. The number of loci and the pedigree size for the analysis described here precluded either of the above algorithms, and necessitated the use of sampling-based approaches.

The idea of a sampling-based approach is that rather than exact summation over possible genotype configurations, genotype configurations are generated at random, and Monte Carlo estimates of probabilities obtained (Thompson, 1994). For example, to get a Monte Carlo estimate of the probability of our observed data Y (genotypic or phenotypic) given some model parameters θ, we could generate N genotype configurations $G_i, i = 1 \ldots N)$ at random and then calculate the Monte Carlo estimate of $p(Y|\theta)$ as:

$$\hat{p}(Y|\theta) = \frac{1}{N}\sum_{i=1}^{N} p(Y|G_i,\theta)\, p(G_i,\theta).$$

This can be inefficient for all but simple problems because typically the vast majority of genotype configurations produced have zero or negligible probability given Y, and so contribute nothing to the summation. A more efficient scheme would be to sample the genotype configurations G_i conditional on Y (and θ). Sampling directly from $p(G|Y,\theta)$ is not possible without knowing $p(Y|\theta)$, the quantity we want to obtain. Markov chain Monte Carlo (MCMC) (Metropolis *et al.*, 1953; Hastings, 1970) methods can be used to produce correlated samples from distributions known only up to a normalizing constant, so can be used to sample from

$$p(G|Y,\theta) \propto p(Y|G,\theta)\, p(G|\theta).$$

The drawback of this approach is that because the genotype samples are correlated, more samples are required to produce a Monte Carlo estimate of a given degree of accuracy than if independent samples were used.

As well as handling missing genotypes, the MCMC approach can also be used in cases where the other model parameters, θ, are also unknown. If we define a joint distribution for all model parameters (treating the genotypes G as parameters), then MCMC can be used to sample from this joint distribution. Estimates of the marginal posterior distribution of any parameter, or group of parameters, of interest can be obtained using this approach. For example, in the context of QTL mapping the model parameters would include the effect and position of the QTL.

(vi) *Model for MCMC analysis*

A MCMC approach was used in this study to perform the linkage analysis of the progressive BC population. The genetic effect of the trait was modelled as due to a single sex-linked QTL and a polygenic effect. Litter was fitted as a random effect (uncorrelated to either of the genetic effects); parity was fitted as a fixed effect; number weaned from a litter and percentage high line genotype were fitted as linear covariates (this last covariate was used to account for the changes in the means and variances of body weight due to the changing percentage of high line autosomal loci). For the QTL, separate effects were estimated for males and females.

Uniform priors were placed on all model parameters and the MCMC sampler was used to integrate out the 'unwanted' parameters. In the case of certain par-

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

ameters, such as QTL position and effects, there could be prior information available about their distribution; using the MCMC framework, this prior information could be easily incorporated into the analysis. Fitting a uniform prior, while not strictly uninformative, imposes little prior constraint on the parameter values. The existence of strongly peaked posterior distributions for the model parameters (as in this study) indicates that there is a lot of information in the data, and therefore the prior is relatively unimportant in this case.

(vii) *MCMC sampling scheme*

A typical MCMC sampling scheme is as follows: All model parameters are set to some initial values such that their joint probability is > 0. Each parameter is updated either singly or in blocks by proposing an update from some known distribution, calculating an acceptance probability A for the update, and then accepting the update with probability A. If the update is not accepted then the parameter value remains unchanged. For the current analysis, two types of update steps were used: Metropolis–Hastings steps (Hastings, 1970) and Gibbs steps (Geman & Geman, 1984). With a Metropolis–Hastings step, the proposal distribution for the update can be almost arbitrary. A Gibbs update step is a special case of this in which a parameter is updated by sampling from its distributional conditional on the current values of all the other model parameters. With a Gibbs update step the acceptance probability A is always 1, so the update is always accepted. More details on these update steps, and their use in genetic analysis, is given in Thompson (1994). The sampling scheme used for the analysis was as follows:

1. update complete marker genotypes (including phase) for each marker locus in turn;
2. update QTL effects;
3. update QTL position;
4. update QTL genotypes;
5. update polygenic effects;
6. update variance components (polygenic and environmental);
7. update fixed effects and covariates;
8. repeat.

All these update steps were Gibbs steps, apart from the updating of QTL position which used a Metropolis–Hastings step, with a proposed new position for the QTL being sampled from a uniform distribution along the chromosome. The genotypes (i.e. at QTL and missing marker genotypes) for all individuals were sampled simultaneously for each locus in turn conditional on the marker observations, the trait values, the model parameters, and the genotypes at other loci. This sampling scheme was suggested by Kong (1991) as a means of improving the mixing characteristics of the sampler, and differs from the sampling schemes typically used in genetics applications, where the genotypes at a given locus are updated on an individual-by-individual basis (e.g. Guo & Thompson, 1992; Heath, 1994). The scheme has some similarities to the blocking Gibbs sampler (Jensen & Kong, 1997) where a portion of the pedigree is updated simultaneously. Sampling of the other model parameters such as the QTL effects, polygenic breeding values, covariate effects and variance components was performed in essentially the same way as described by Wang *et al.* (1993) and Heath (1994). The procedure was extensively tested on simulated data (Heath *et al.*, 1997).

(viii) *Estimating posterior distributions*

Estimates of the posterior marginal distribution of any parameter of interest can be obtained from the sequence of realizations for that parameter produced by the MCMC scheme. For the analysis presented in this paper, estimates were obtained of the posterior densities of the QTL positions and effects. The confidence intervals for the QTL positions were obtained from the ± 2 SE of the mean position as proposed by Darvasi *et al.* (1993).

3. Results

(i) *Changes in the distribution of traits measured over the generations of backcrossing*

Means and variances of the body weight traits measured in the BC population, corrected for the fixed effects and covariates, are shown for each generation in Table 3. Mean 6 week and 10 week body weights tended to decline with each generation of backcrossing to the inbred low line, with the exception of generation 2 for which 6 week weights were available only on females. The reduction in mean body weights were also associated with a reduction in the variance of the body weight measurement. This reduction is expected, as the percentage contribution of high line autosomal linked QTLs for body weight decline by an average of 50% with each generation of backcrossing. In the BC5 generation, individuals carry on average 1·56% high line autosomal alleles.

(ii) *Effects associated with chromosome segments*

An analysis was carried out to estimate the effects associated with carrying specific segments of the high line X chromosome on the inbred low line background. The segment effects estimates (Table 4) represent the difference between the two hemizygous genotypes in males ($2a$) and the difference between the homozygous low and heterozygous genotypes in females ($a+d$). In males, the marker segments at the most distal end of the X chromosome (between *DXMit16* and *DXMit31*, segments 1–3) are associated with small or negative differences in body weight (these differences are not

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

Table 4. *Mean effect* (*g*) *for 10 week weight associated with segments of the X chromosome in males and females from generations 3, 4 and 5 of the progressive backcross population* (*standard errors, SE, shown in parentheses*)

Estimated mean effect (g)				Segment	*DXMit* marker								
Males	(SE)	Females	(SE)	no.	55	187	50	25	16	79	38	121	31
0·48	(0·89)	−2·74	(2·00)	1								✓	✓
−0·15	(0·96)	−1·05	(0·59)	2						✓	✓	✓	
−0·38	(0·72)	−1·80	(0·91)	3					✓	✓	✓		
4·26	(0·95)	2·03	(0·67)	4			✓	✓	✓				
4·41	(0·75)	1·82	(0·59)	5	✓	✓	✓	✓					
3·81	(0·57)	1·09	(0·42)	6	✓	✓	✓						
1·11	(0·83)	−0·12	(0·77)	8	✓	✓							
5·50	(0·84)	1·65	(0·78)	9	✓	✓	✓	✓	✓	✓	✓		

The ticks represent markers which can be traced unequivocally to the high selection line. The remaining markers on the X chromosome can be traced to the low inbred line.

significantly different from zero, $P > 0{\cdot}05$). The estimated effects associated with the segments containing markers at the most proximal end of the X chromosome (between *DXMit55* and *DXMit16*, segments 4, 5, 8 and 9) are significantly different from zero ($P < 0{\cdot}05$), and explain approximately 4 g of the difference in mean body weight between the two hemizygote segment groups in males. The estimated effect associated with marker segment 7 (*DXMit55* to *DXMit187*) was not significantly different from zero, however. The same trends are present in the females from the BC generations 3, 4 and 5, although the estimated effects were approximately half of the estimates in males. Similar trends can be seen in the estimates associated with segments for 6 week body weight (data not shown). It should be noted that the origin of the segments of X chromosome between high and low line markers (i.e. where there has been a recombination event) are not precisely known, but with each generation of backcrossing the percentage of high line X chromosome between these markers will tend to decline. The results obtained from the segment analysis show that the distal end of the X chromosome (*DXMit16* to *DXMit31*) is not associated with a QTL for body weight, but there is a significant association between the proximal end of the X chromosome and body weight at 6 and 10 weeks (between markers *DXMit50* to *DXMit25*).

(iii) *MCMC sampling*

Fig. 2 show the estimated marginal posterior distribution of the QTL position, using a uniform prior over the whole X chromosome for the QTL position. The distribution was obtained from the frequency at which the QTL was estimated as being at each position along the chromosome. The plot shows only the region of the X chromosome of 21 cM to 32 cM, as the probability of the QTL in all sections outside this area was effectively zero. The mean estimated QTL position was 25·4 cM with an estimated 95% confidence interval of $\pm 2{\cdot}8$ cM for 10 week weight, and $26{\cdot}9 \pm 3{\cdot}0$ cM for 6 week weight. The analysis therefore points to the QTL position being within the marker interval *DXMit50* (19·8 cM) to *DXMit*25 (33·7 cM). The irregular shapes of the curves in Fig. 2 are due to Monte Carlo (sampling) error.

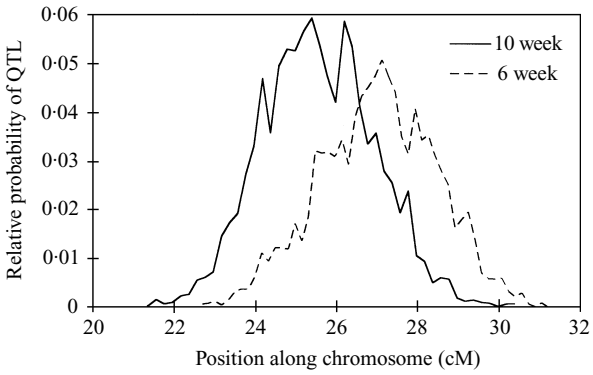

Fig. 2. Marginal posterior distributions for QTL positions, obtained from the MCMC analysis. The nearest flanking markers are *DXMit50* and *DXMit25* at 19·8 and 33·7 cM respectively.

Table 5. *Estimated QTL effects for 6 week and 10 week weight obtained from the MCMC analysis, shown as deviations from the hemizygous low line genotype in males* (*a*), *and homozygote low genotype in females* (*aa*)

	Estimated QTL effect (g)	
Trait	Males	Females
6 week weight	3·4 (0·2)	1·4 (0·2)
10 week weight	4·3 (0·2)	1·8 (0·3)

Standard errors of estimates are shown in parentheses.

The mean QTL effects estimated using the MCMC method (when the QTL effects were sampled at the mean estimated QTL position) are presented in Table 5. The estimated marginal posterior distribution for

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

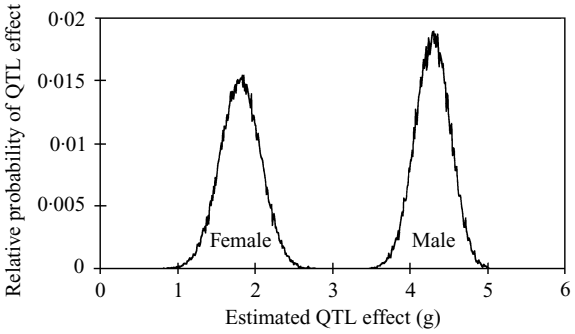

Fig. 3. Marginal posterior distributions for QTL effects on 10 week weight, shown as the difference between the hemizygotes in males and the difference between the homozygous low and heterozygous genotypes in females.

QTL effects (Fig. 3) shows that the sampled QTL effects are approximately normally distributed. The mean QTL effects are similar in males and females. If we assume no dominance in the females, the estimated QTL effects (representing the difference between the hemizygous males and 2(Aa − aa) in females) were approximately 3 g at 6 weeks and 4 g at 10 weeks in both males and females. These estimated QTL effects are close to those obtained using the segment analysis (Table 4).

4. Discussion

(i) *Effects of marked X chromosome segments*

The estimated body weight effects associated with segments at the most distal end of the X chromosome (those segments containing markers *DXMit79* to *DXMit31*) are not significantly different from zero in both males and females. However, segments at the proximal end of the X chromosome are associated with estimated effects of approximately 4 g in males and between 1 g and 2 g in females. Segment 8 (Table 4) is not associated with a significant effect, implying that the QTL is not in this segment. Segments 4 and 5 overlap between markers *DXMit50* and *DXMit25*, plus an unknown amount flanking them. Taken together with the significant effects associated with segments 6 and 9, the data are consistent with a QTL in the marker interval *DXMit50* to *DXMit25*, or immediately flanking these markers, with a total effect (2*a*) of approximately 4 g in males and females, assuming no dominance in females. The results do not allow the conclusion that the X-linked QTL is in the interval *DXMit50* to *DXMit25*, as the points at which the recombination break points occurred in the intervals *DXMit187* to *DXMit50* and *DXMit25* to *DXMit16* are not known. The position of the QTL can be estimated more precisely by using information from all individuals in the BC pedigree with the MCMC sampling-based method.

(ii) *MCMC analysis of the BC population*

The results presented in Fig. 2 show the estimated marginal posterior distribution of the QTL position on the X chromosome for 6 week and 10 week weight data. The frequency at which a QTL was sampled outside the interval 21 cM to 32 cM was close to zero, therefore strongly indicating that the QTL is positioned in the marker interval *DXMit50* to *DXMit25*, with no evidence for a second QTL elsewhere on the X chromosome. The estimated positions of the QTL (±95% CI) for 6 week weight and 10 week weight within this interval were 25·4 ± 2·8 cM and 26·9 ± 3·0 cM respectively. The positions estimated for the two body weight measurements therefore show overlapping CI for the QTL position.

(iii) *Combining information from the F_2 study*

Combining the estimated QTL positions obtained from the F_2 population described in the associated study (Rance *et al*., 1997) and the present study points to a single X-linked QTL (or a number of very closely linked QTLs) positioned in the marker interval

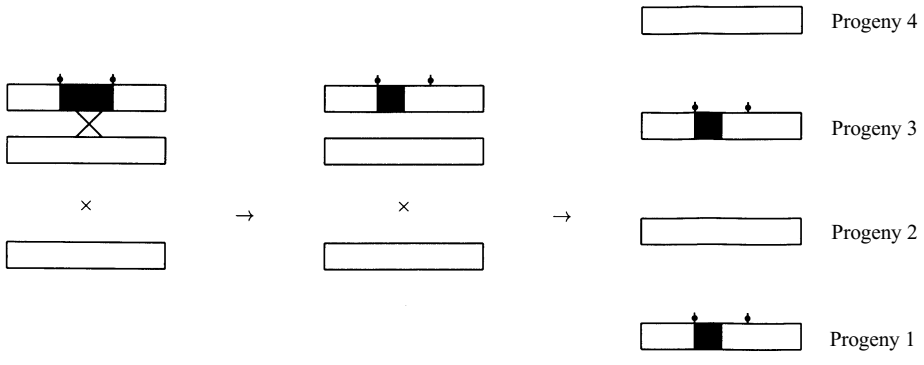

Fig. 4. Crossing scheme for progeny test with female recombinants.

https://doi.org/10.1017/S0016672397002929 Published online by Cambridge University Press

DXMit50 to *DXMit25*. Pooling the estimates for the QTL positions obtained in the two populations by weighting the estimates by the reciprocal of their sampling variance, the estimated QTL position (±SE) was 24·6±1·2 cM for 10 week weight and 25·8±1·3 cM for 6 week weight. It should be noted that the QTL positions were estimated using two different methods in the two experiments: by multiple regression in the F_2 population, and by MCMC sampling assuming uniform priors for the QTL parameters in the progressive backcross population. When the estimates are expressed as a percentage of the mean body weight, the QTL appears to have a very similar percentage contribution in both populations, which suggests the X-linked QTL has a multiplicative effect on mean body weight. The single X-linked QTL appears to explain almost the whole contribution of the X chromosome, increasing mean body weight by approximately 18%.

(iv) *Strategy for positional cloning*

QTL mapping is necessarily a multi-stage process, with a logical end point of map-based cloning, but so far there have been no reports of the positional cloning of QTLs. It is becoming apparent that the effort involved in fine-scale mapping of a QTL exceeds that for an initial screen of the entire genome. Mapping to a sub-centimorgan level is required to positionally clone any gene, and ultimately depends on knowing the phenotype of individuals containing rare recombination events. In fine-scale mapping of QTLs, an additional problem is the lack of a one-to-one correspondence between a recombinant individual's phenotypic value for the trait and its QTL 'genotype', so each recombinant provides less information than for mapping a Mendelian locus. As proposed by Lander & Botstein (1989), 'progeny testing' of individuals known to be recombinant between markers flanking a region containing a QTL can allow the same resolution as for mapping a Mendelian locus.

A progeny testing scheme suitable for an X-linked QTL is shown in Fig. 4. A suitable starting point for the cross would be a congenic line containing high line markers *DXMit50* and *DXMit25*, which flank the QTL, on the low line inbred background. Females heterozygous for the segment would be crossed with low line inbred males to generate progeny recombinant between the two flanking microsatellite markers, detected by genotyping at the flanking markers. Such recombinant females would then be crossed to inbred low line males, causing the recombinant segment to segregate in their male progeny. If the QTL is in the recombinant segment, a large difference in body weight would be observed between individuals containing or not containing the segment. The exact position of the recombination break point in the recombinant individual, base on further markers, would be the information required to fine-scale map the QTL, and to allow it to be cloned by the same strategies as for a Mendelian locus.

References

Cannings, C., Thompson, E. A. & Skolnick, M. H. (1978). Probability functions on complex pedigrees. *Advances in Applied Probability* **10**, 26–61.

Darvasi, A., Weinreb, A., Minke, V., Weller, J. I. & Soller, M. (1993). Detecting marker–QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* **134**, 943–951.

Dietrich, W., Katz, H., Lincoln, S. E., Shin, H., Friedman, J., Dracopoli, N. C. & Lander, E. S. (1992). A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**, 423–447.

Elston, R. C. & Stewart, J. (1971). A general model for the genetic analysis of pedigree data. *Human Heredity* **21**, 523–542.

Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

Genstat 5 Committee (1993). *Genstat 5 Release 3 Reference Manual*. Oxford: Clarendon Press.

Guo, S. W. & Thompson, E. A. (1992). A Monte Carlo method for combined segregation and linkage analysis. *American Journal of Human Genetics* **51**, 1111–1126.

Haley, C. S. & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315–324.

Hastings, I. M. & Hill, W. G. (1989). A note on the effect of different selection criteria on carcass composition in mice. *Animal Production* **48**, 229–233.

Hastings, I. M. & Veerkamp, R. F. (1993). The genetic basis of response in mouse lines divergently selected for body weight or fat content. I. The relative contributions of auto-somal and sex-linked genes. *Genetical Research* **62**, 169–175.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109.

Heath, S. C. (1994). Estimation of linked QTL effects with an animal model using Gibbs sampling. In *Proceedings of the Fifth World Congress on Genetics Applied to Livestock Production*, vol. **18**, pp. 398–401.

Heath, S. C., Snow, G. L., Thompson, E. A., Tseng, C. & Wijsman, E. M. (1997). MCMC segregation and linkage analysis. Proceedings of Genetic Analysis Workshop 10. *Genetic Epidemiology* (in press).

Jansen, R. C. (1993). Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205–211.

Jansen, R. C. (1996). A general Monte Carlo method for mapping multiple quantitative trait loci. *Genetics* **142**, 305–311.

Jensen, C. S. & Kong, A. (1997). Blocking Gibbs sampling for linkage analysis in large pedigrees with many loops. Technical Report R96–2048. Dept. Computer Science, Aalborg University, Denmark.

Kong, A. (1991). Analysis of pedigree data using methods combining peeling and Gibbs sampling. *Computer Science and Statistics. Proceedings of the 23rd Symposium on the Interface*, pp. 379–385.

Lander, E. S. & Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199.

Lander, E. S. & Green, P. (1987). Construction of multilocus genetic maps in humans. *Proceedings of the National Academy of Sciences of the USA* **84**, 2363–2367.

Martinez, O. & Curnow, R. N. (1992). Estimating the

locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics* **85**, 480–488.

Mbaga, S. H. (1996). Analysis and inferences from long-term quantitative genetic selection experiments. PhD thesis, University of Edinburgh.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, W. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1091.

Rance, K. A., Hastings, I. M., Hill, W. G. & Keightley, P. D. (1994). Mapping of putative QTL influencing body weight on the X chromosome of mice. In *Proceedings of the Fifth World Congress on Genetics Applied to Livestock Production*, vol. 21, pp. 268–269. University of Guelph.

Rance, K. A., Hill, W. G. & Keightley, P. D. (1997). Mapping quantitative trait loci for body weight on the X chromosome in mice. I. Analysis of a reciprocal F_2 population. *Genetical Research* **70**, 117–124.

Routman, E. & Cheverud, J. (1994). A rapid method of scoring simple sequence repeat polymorphisms with agarose-gel electrophoresis. *Mammalian Genome* **5**, 187–188.

Sharp, G. L., Hill, W. G. & Robertson, A. (1984). Effects of selection on growth, body composition and food intake in mice. 1. Responses in selected traits. *Genetical Research* **43**, 75–92.

Shrimpton, A. E. & Robertson, A. (1988*a*). The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. I. Allocation of third chromosome sterno-pleural bristle effects to chromosome sections. *Genetics* **118**, 437–443.

Shrimpton, A. E. & Robertson, A. (1988*b*). The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. II. Distribution of third chromosome bristle effects within chromosome sections. *Genetics* **118**, 445–459.

Thompson, E. A. (1994). Monte Carlo likelihood in genetic mapping. Statistical Science **9**, 355–366.

van Ooijen, J. W. (1992). Accuracy of mapping quantitative trait loci in autogamous species. *Theoretical and Applied Genetics* **84**, 803–811.

Veerkamp, R. F., Haley, C. S., Knott, S. A. & Hastings, I. M.(1993). The genetic basis of response in mouse lines divergently selected for body weight or fat content. II. The contribution of genes with a large effect. *Genetical Research* **62**, 177–182.

Wang, C. S., Rutledge, J. J. & Gianola, D. (1993). Marginal inferences about variance components in a mixed linear model using Gibbs sampling. *Genetics*, *Selection*, *Evolution* **25**, 41–62.

Zeng, Z.-B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468.