INTERVAL MAPPING

C.S. Haley and S.A. Knott
Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, U.K.
ICAPB, University of Edinburgh, West Mains Rd., Edinburgh, EH9 3JT, U.K.

SUMMARY

In interval mapping pairs or groups of marker loci are used for the identification of quantitative trait loci (QTLs) lying between those markers. With data from a segregating population from an inbred line cross the markers flanking each interval can be considered a pair at a time. The method can be applied by regression or maximum likelihood (ML), with the two approaches giving similar results. Regression is simpler and hence allows the fitting of more realistic models. Interval mapping gives some extra power to detect QTLs and better separation of the estimates of the size and location of the QTLs compared to using markers one at a time. Groups of linked QTLs can mislead the experimenter if a model with only a single QTL is fitted to the data, but models fitting multiple QTLs or a single QTL and additional markers as cofactors can ameliorate this problem. In outbred populations marker loci may not be fully informative and this will reduce the power to detect a QTL and may lead to biased estimates of its position. For both the analysis of crosses between outbred lines and the analysis of half-sib structured outbred populations there are simple ways to combine information from a number of markers in a linkage group to increase the power to detect QTLs and reduce biases in estimates of QTL position and effect. The benefits of using multiple, as opposed to single, markers are greater in the analysis of outbred populations than they are in the analysis of inbred line crosses. Future experimental studies of QTLs in livestock populations should attempt to use multiple marker methods in order to reap these benefits.

INTRODUCTION

Interval mapping is concerned with the identification and localisation on a map of the genome of individual loci (or clusters of linked loci) that contribute to variation in quantitative traits. There are three main justifications for the mapping of such quantitative trait loci (or QTLs). First, to examine the failings of the infinitesimal model of quantitative genetic variation and perhaps improve it. Second, to allow marker assisted selection to manipulate loci within and between breeds. Third, ultimately to clone some of the loci involved and study (and possibly manipulate) their actions at the molecular level.

In early QTL mapping studies markers were used one at a time, however the development of complete marker maps led to the realisation that it was possible to use pairs or groups of linked loci for mapping QTLs. The term 'interval mapping' was coined to describe the mapping of a QTL between a pair of linked markers (Lander and Botstein, 1986, 1989), although the use of multiple loci for mapping is an extension of an approach used for marker linkage studies (e.g. Ott, 1991). Interval mapping has largely been used in the analysis of data from inbred line crosses. Such populations illustrate the basic features of interval mapping and we open by discussing these. Inbred lines are rarely or never available for livestock, however, so we will move on to the application of the principles of interval mapping to crosses between outbred lines and to random mating populations.

INTERVAL MAPPING - INBRED LINES

Basic approach

The basis of any mapping of QTLs is an association between phenotypic variation and segregation at one or more genetic markers. For any gene, marker or QTL, to be segregating in a population (e.g. F_2 or backcross) derived from an inbred line cross it must have been fixed for alternative alleles in the two inbred lines crossed and so have a heterozygosity of one in the F_1 population. Hence codominant loci (as most DNA based markers are) in such a cross are fully informative in the sense that it is possible to identify unequivocally (barring genotyping errors, missing data, etc.) whether individuals in a segregating generation have none, one or two alleles from each inbred line. Thus for a putative QTL at a given recombinational distance from a single marker it is simple to calculate the probability of individuals in the segregating generation having each possible QTL genotype conditional on their marker genotype. However, the expected mean difference between marker genotypes due to the linked

QTL is a function of both the effect of the QTL and its distance from the marker, so with information on the mean difference between genotypes at a marker, it is not possible to estimate both the effect of the QTL and its distance from the marker. Segregation of the QTL within a marker genotype due to recombination will affect the within marker genotype phenotypic distribution. In principle, by using maximum likelihood (ML) or similar analyses this information can be combined with that from mean differences to estimate separately the effect of a QTL and its distance from the marker (Weller, 1986; Lander and Botstein, 1989; Haley, 1991; Knott and Haley, 1992a). However, in practice such estimates may be rather poor (see below).

Interval mapping is a simple extension of single marker mapping possible when a map of linked markers is available. For a putative QTL at a given position between two markers, the probability of individuals in the segregating generation being each possible QTL genotype is calculated conditional on the genotype at both markers. The number of marker classes is increased compared to using a single marker (e.g. from 2 to 4 in a backcross) and, for the same distance from a marker, the probability with which we can assign putative QTL genotypes is increased. Furthermore, there will only be one combination of QTL effect and position between the markers which best explains the mean phenotypic differences between the marker genotype classes. Thus the ability to separately estimate the QTL effect and position is greatly improved. For fully informative markers with no interference, the markers flanking an interval absorb the effect of a QTL between them - there is no advantage to be gained by the inclusion of more than the two flanking markers.

Lander and Botstein (1989) applied the method by ML, maximising the likelihood for defined positions along the chromosome (e.g. 1 cM intervals). Various formulations of the likelihood for one marker or for interval mapping have been given (e.g. Weller, 1986; Lander and Botstein, 1989; Knott and Haley, 1992a; Jansen, 1992). Evidence for a QTL at a particular position is obtained as the logarithm of the ratio of this likelihood to that which assumes no QTL. This test statistic is then plotted against the position on the chromosome to obtain a curve, the highest point of which indicates the most likely position of a QTL. Lander and Botstein (1989) plotted $\log_{10}$ of this ratio, obtaining the LOD (logarithm of the odds) score familiar to human geneticists, using $2\log_e$ of this ratio (i.e. $\approx 4.6 \times$ LOD) gives the (equivalent) likelihood ratio test statistic more familiar to animal breeders. In this paper, to ease comparisons, we have converted results originally reported as LODs to likelihood ratio test statistics.

Advantages of interval mapping

The original claims for interval mapping over single marker mapping were that it allowed better separation of position and effects of a QTL, required fewer progeny to detect a given effect (9 to 28% fewer for a backcross with a QTL at the midpoint between two markers 10 to 40 cM apart) and could be used to distinguish a pair of linked QTLs from a single QTL (Lander and Botstein, 1989). As far as differences in power between the methods are concerned, subsequent research have found these to be less than originally claimed. Lander and Botstein (1989) made a comparison between a test based on one marker and interval mapping. In practice the difference in power between the methods is reduced because both single markers flanking an interval would be tested individually and the power would be

Table 1. Mean estimates and their empirical standard deviation from interval and single marker mapping applied by ML (Knott and Haley, 1992a). Results based upon 50 replicate simulations of 1000 F_2 individuals with a QTL placed at 25 cM on a 100 cM chromosome with markers 10 or 50 cM apart.

	Simulated effect		Interval mapping		Single marker mapping	
Parameter	of QTL[1]	Marker spacing:	10 cM	50 cM	10 cM	50 cM
Additive effect (a)[1]	0.25		0.260	0.250	0.533	0.437
Empirical s.d. of a			0.047	0.063	0.269	0.293
Additive effect (a)	0.5		0.500	0.497	0.624	0.601
Empirical s.d. of a			0.047	0.061	0.183	0.256

[1]Represents half the difference between homozygotes in residual standard deviations.

dictated by that giving the highest test statistic (Knott and Haley, 1992a; Darvasi *et al.*, 1993). Furthermore, Lander and Botstein (1989) assumed the QTL lay midway between the markers and at this point there is maximum benefit from using interval mapping (Darvasi *et al.*, 1993). Interval mapping is better at separating the effect of a QTL from its position on the map, giving less biased estimates with much lower empirical standard errors than single marker mapping (Table 1). Further advantages of interval mapping are that it is more robust than single marker mapping to non-normality in the data and hence is not biased by the presence of unlinked QTLs (Knott and Haley, 1992a). The presence of linked QTLs outside the interval under study can cause problems, however (see below).

ML versus least squares

Most applications of interval mapping have been by ML, however it is also possible to apply the method by regression and this has some advantages (Haley and Knott, 1992; Martinez and Curnow, 1992). In the regression approach the probabilities of genotypes at a QTL at a given position between two markers are calculated conditional upon the marker genotypes, just as they are for ML. Then the expected average additive and dominance deviations of individuals are calculated from these probabilities (e.g. $C_a = p_{QQ} - p_{qq}$ and $C_d = p_{Qq}$, where C_a and C_d are the expected coefficients of the additive and dominance deviations, respectively, and p_{QQ}, p_{Qq} and p_{qq} are the conditional probabilities of being QTL genotypes QQ, Qq and qq, respectively). Regressing the phenotypic values onto these coefficients provides estimates of a and d for that position. Repeating this procedure at intervals (e.g. 1 cM) along the chromosome indicates the most likely position of a QTL as that at which the regression variance (F) ratio is maximised. A likelihood ratio test statistic from the regression approach is $[N \log_e(\text{RSS(model omitting QTL)/RSS(model including QTL)})]$, where there are N observations. Thus estimates and test statistics from the ML and regression approaches can be directly compared (Table 2). As can be seen from Table 2, ML and regression applications of interval mapping produce very similar estimates and test statistics which are highly correlated over replicates. The regression approach does not utilise information on the phenotypic distribution within marker classes. Thus the similarity of results from the two methods indicates that the great majority of information for estimating QTL effects comes from the mean differences between marker classes. As the two methods produce such similar results and the regression method is simple and easily extended to include the effects of multiple QTLs and fixed effects, it will be the analytical method of choice in many situations.

Table 2. Comparison of estimates from regression and ML applications of interval mapping (Haley and Knott, 1992). A QTL of effect $a = 0.25$ residual standard deviations was simulated at position 25 cM on a 100 cM chromosome. Fifty replicates of 1000 F_2 individuals were analysed.

	10 cM spaced markers				50 cM spaced markers			
	Effect (a)	Position (cM)	Residual s.d.	Test statistic	Effect (a)	Position (cM)	Residual s.d.	Test statistic
Mean estimate - ML	0.258	25.6	0.997	33.1	0.247	24.2	0.996	20.5
Empirical s.d. - ML	0.047	3.0	0.020	11.8	0.069	10.6	0.020	8.8
Mean estimate - Regression	0.258	25.6	0.999	33.2	0.247	24.0	1.001	20.5
Empirical s.d. - Regression	0.048	3.1	0.020	12.0	0.068	10.6	0.020	8.8
Correlation between methods	0.999	0.990	0.999	1.000	0.997	0.997	0.976	0.999

Accounting for multiple QTLs

Lander and Botstein (1989) suggested that multiple linked QTLs would result in distinct separate peaks in the QTL likelihood surface. They suggested that in this case a QTL should be fitted at the position of the higher peak and then the best position for a second QTL should be found by addition to this model. If linked QTLs are of large effect and well separated (>50 cM) on a relatively marker dense chromosome they may be detected as separate peaks on the likelihood curve. However this will often not be the case and the effect of two linked QTLs can combine to produce a single peak at a position between them, leading to biased estimates of QTL position and effect (Haley and Knott, 1992; Knott and Haley, 1992a; Martinez and Curnow, 1992). One solution to the problem is to fit two QTLs

simultaneously, exploring all possible positions on the chromosome in a two dimensional grid. Computational considerations make this much easier to achieve using regression than using ML. Results from simulations show that good estimates of the positions and effects of two linked QTLs (even in the presence of interactions between them) at 50 cM separation can be obtained (Haley and Knott, 1992).

Extension of the method above to the mapping of multiple QTLs would be intractable and methods have been proposed to circumvent this problem (Jansen, 1993; Zeng, 1993). In these methods, whilst using interval mapping in one interval, additional markers are included as cofactors in the analysis. This has two potential benefits. First, it reduces the bias due to the effects of QTLs outside, but linked to, the interval under study. Second, by removing the effects of unlinked QTLs it can reduce the residual variance and hence increase the power to detect a QTL in the interval under study. With many markers it is not possible to fit them all as cofactors and so Jansen (1993) suggests selecting the most important markers by regression using backward elimination. Whilst these approaches have yet to be fully explored, they show substantial promise for removing the effects of linked and unlinked QTLs in a computationally feasible way (Jansen, 1993).

Significance thresholds

Interval mapping involves testing for the presence of a QTL at multiple positions (e.g. 1 cM intervals) through the genome, thus if the usual 5% threshold for the likelihood ratio test (in a ML analysis) or F ratio (in a regression analysis) is used type I errors will be a problem (i.e. there will be too many false positive results). The significance threshold will depend upon both the size of the genome under study and the density of the markers. Lander and Botstein (1989) have derived approximations to predict the threshold for a chosen whole-genome level of false positive results in the analysis of a backcross when no QTLs are present. Some examples are given in Table 3. Note, however, that the threshold has to be increased slightly when analysing an F_2 population and so estimating both an additive and dominance effect of the QTL (van Ooijen, 1992). The threshold required in more complex models (e.g. when fitting additional markers in the model) has yet to be explored.

Table 3. Likelihood ratio significance thresholds for a whole genome type I error of 5% in an analysis of a backcross by interval mapping. Based on Lander and Botstein (1989) and Darvasi *et al.* (1993).

	Distance between markers (cM)			
	0	10	20	50
'Pig' (18 autosomes, total map length 18 Morgans)	15.2	12.4	12.0	10.6
'Cattle' (29 autosomes, total map length 29 Morgans)	16.5	13.3	12.8	11.4

A further point concerns the null hypothesis used - that there are no QTLs segregating in the population. This is obviously wrong when the lines crossed had different means. We have previously shown that the combined effect of a large number of QTLs of small effect will be to inflate the test statistic over the whole genome (Knott and Haley, 1992a). So the test statistic may be above the significance threshold in several places due to chance even if no QTLs of large effect are segregating. Fitting markers outside the interval under study (e.g. Jansen, 1993; Zeng, 1993) may ameliorate this problem by reducing the extent to which the test statistic is inflated by multiple linked QTLs.

Power

It is important to know whether a given study has the power to detect QTLs if they are present. Power is a function of two factors - the expected test statistic produced by a QTL of any particular effect and the significance threshold appropriate for the study and chosen method of analysis. The expected test statistic for a single QTL unlinked to any others can be adequately predicted by formula provided by Lander and Botstein (1989, see also Knott and Haley, 1992a)) for the expected likelihood ratio (ELRT):

1) $ELRT = [(1 - 2\theta)/(1 - \theta)]\, N \log_e[1/(1-p)]$ 2) $ELRT = N \log_e[1/(1-p)]$

where 1) applies to a QTL midway between two markers a recombination fraction of θ apart and 2) to a

QTL at the position of a marker (the two extreme cases) and the QTL explains a proportion p of the total variance in a population of size N. Note that ELRT is proportional to the sample size and approximately proportional to the variance due to the QTL. These ELRT predictions can be combined with the thresholds in Table 3 above to calculate the QTL effect needed for 50% power of detection (Table 4).

Table 4. Size of QTL effect (% of total variance) for 50% power of detection in a backcross mapping study. The worst case (QTL midway between two markers) is given using 'pig' thresholds from Table 3.

	Distance between markers (cM)			
Population size	0	10	20	50
200	7.3	6.7	7.2	9.4
500	3.0	2.7	2.9	3.9

These size of effect detectable does not increase greatly as marker spacing decreases because the significance threshold increases almost commensurately with the test statistic. Thus decreasing the marker spacing below 20 cM does little to increase the power (see also Darvasi *et al.*; 1993). One final point to note is that the power of methods based upon ML has been found to be very similar to that based upon least squares or regression (Haley, 1991; Haley and Knott, 1992; Simpson, 1992; Darvasi *et al.*, 1993).

QTL mapping accuracy

An important issue, both for marker assisted selection and for positional cloning of QTLs, is the accuracy with which QTLs can be mapped. Simulation studies suggest that this accuracy may not be very great (Table 5). Even the most accurately mapped QTLs would represent a huge length of DNA to explore (assume 4 x 1.9 cM represents the 95% confidence interval of position, this would be about 8 x 10^6 DNA base pairs). Furthermore, even QTL of large effect may often be placed in the wrong interval when markers are close together. See also Darvasi *et al.* (1993).

Table 5. Empirical standard deviation on estimates of QTL position and number of estimates in the correct interval. Based upon 50 replicates of 1000 F_2 individuals with a QTL placed at 25 cM on a 100 cM chromosome (Knott and Haley, 1992a).

Simulated		Empirical standard deviation		Correct interval (out of 50)	
effect of QTL[1]	Marker spacing:	10 cM	50 cM	10 cM	50 cM
0.25		7.5	14.3	36	47
0.5		1.9	4.8	49	50

[1]Represents half the difference between homozygotes in residual standard deviations

A related problem is the confidence interval on a single estimate of position. Lander and Botstein (1989) use a one LOD support interval (the region of chromosome over which $\log_{10}$ of the likelihood ratio is no more than one less than the maximum LOD) to provide some estimate of the confidence interval. By analogy with linkage studies (Ott, 1991), a one LOD support interval should be equivalent to a 95% confidence interval. Simulation, however, suggests that a two LOD support interval is closer to a 95% confidence interval (van Ooijen, 1992). Finally, all these studies have been under ideal conditions, with only a single QTL on the chromosome and no other disturbing influences. As such factors may be present in real data, estimates of QTL position should be treated with some caution.

Conclusions

Interval mapping does have some real advantages for localising QTLs in populations derived by crossing inbred lines. However, problems remain, such as multiple linked QTLs, and optimum strategies for applying interval mapping are only just being developed. For detecting QTLs, dense marker maps are not required, although they can improve the accuracy with which QTLs are mapped.

INTERVAL MAPPING - LIVESTOCK POPULATIONS

Introduction

It is much more difficult to map QTLs in livestock populations than it is using inbred line crosses. Problems include a lower heterozygosity of both markers and QTLs, exacerbated by the fact that within populations QTLs of large effect may have been fixed by selection, marker-QTL linkage phase will differ between families, pedigrees are typically complex and between family genetic variation is present and can bias tests for QTLs. All of these factors reduce power and have to be accounted for in analyses.

ML approach

We have previously shown how ML may be used to apply interval mapping to the analysis of data from a population of nuclear full-sib families (Knott and Haley, 1992b), and the method could be easily adapted for other structures such as half-sibs. The likelihood was developed assuming Hardy-Weinberg equilibrium at the markers and QTLs and linkage equilibrium between the loci. A random component between families was also included to account for effect of unlinked genes or between family environmental effects and numerical integration over this component made the method computationally tractable. Based on the application of the method to simulated data we found that reasonable estimates of the simulated effects could be obtained. However, omission of the between family component from the model led to an overestimation of the QTL effect - the QTL absorbed some of the between family variance left unexplained by the model. As expected, for a given number of full-sib progeny, the test statistic increased with family size and also with the heterozygosity at the marker loci. The relative advantage of flanking markers over single markers was greater than it was for data from inbred line crosses (Table 6). Finally, because the test statistic was greatly influenced by the heterozygosity of the markers, the highest test statistic may be contained in an interval flanked by more informative markers, rather than the interval containing the QTL - this would lead to biased estimates of QTL position and effect. So although the ML method of interval mapping can be applied to data from outbred populations and produce reasonable results, there are drawbacks, not the least being that the methods are relatively computationally demanding.

Table 6. Mean likelihood ratio test statistics from the analysis of full-sib data. Based on 10 replicates of 1000 sibs with their parents, with markers of high (8 alleles) or low (2 alleles) information content and two sizes of interval. The additive QTL had two alleles at equal frequency with two standard deviations between homozygotes, which would produce a test statistic of 320 in a 20 cM interval in 1000 F_2 individuals (Knott and Haley, 1992b).

Size of sibships	No. of marker alleles	Interval size (cM)	Single marker test statistic	Interval mapping test statistic
20	8	20	63.2	84.8
10	8	20	44.5	57.5
10	2	20	18.1	26.7
10	8	50	16.2	21.9
10	2	50	7.4	9.5
4	8	20	17.5	19.9

Regression methods for multiple markers

Many of the methods proposed for the analysis of data from livestock have been based upon least squares (e.g. Niemann-Sorensen and Robertson, 1961; Weller *et al.*, 1990). These methods use markers singly and thus may suffer from loss of information and potentially from biased estimates (the position and the effect of the QTL are confounded and the marker with the most significant effect may be the most informative rather than the closest). We have therefore explored the use of regression for combining information from multiple markers in the analysis of data from outbred populations.

Populations derived from crosses between outbred lines (e.g. breeds or selection lines) are possible

for all livestock species. These resemble inbred line crosses in that it may often be reasonable to assume that QTLs with the largest effects will be fixed for alternative alleles in the two lines. Many of the markers, however, will be segregating within the lines and so they will not be completely heterozygous, and hence informative, in the F_1 cross between the lines. We have recently developed a regression method which uses information from multiple markers in this situation (Haley *et al.* 1994). This method is an extension of the method used for inbred lines in that, on the assumption that the lines crossed were homozygous for alternative QTL alleles, the genotypes at the putative QTL in the segregating progeny (e.g. F_2 or backcross) at a given position in the genome are predicted conditional on the marker genotypes. For inbred lines the markers used are the two that flank the chosen position of the QTL. In data from outbred line crosses the markers that are informative may differ between the two parents (e.g. F_1's) of the segregating generation and this means that up to four markers (two flanking the QTL position in each parent) may be used to calculate the conditional probabilities for each individual. The markers used will differ between families, however, and so over the whole population it may be necessary to use information from all the markers in a linkage group. Subsequently, the phenotypic values are regressed onto the additive and dominance coefficients calculated from the conditional probabilities of QTL genotypes to obtain estimates of these effects at the chosen position. The position at which the test statistic (likelihood ratio or regression F ratio) is maximised is the most likely position for any QTL.

This method has been applied to simulated data and compared with results of analyses which use only information from markers immediately flanking the interval containing the QTL (Table 7). When markers were not completely informative, using all markers increased the test statistic and reduced the empirical standard deviation of parameter estimates compared to using only flanking markers. The advantage of using a relatively dense marker map is greater when markers are not completely informative as with a dense map there is a greater chance that information missing from a marker flanking the position of a QTL can be replaced by information from another nearby marker. We also found that the method could compensate for variation in marker information content along the chromosome and hence remove biases in the estimated position and effect of a QTL that occurred when the only flanking markers were used. This method has recently been used for the detection of QTLs affecting growth and fatness traits in a Wild Boar x Large White cross (Andersson *et al.*, 1994).

Table 7. Mean likelihood ratio test statistic from the analysis of an outbred line cross. Based on 100 replicates of 500 individuals, with three types of marker and three sizes of interval between markers. The additive QTL was at 30 cM on a 100 cM chromosome and was fixed for alternative alleles in the two lines with one standard deviation between homozygotes (Haley *et al.*, 1994).

Marker type	Interval size (cM)	Mean test statistic		Empirical s.d. of position	
		Flanking	All	Flanking	All
Lines fixed for alternative alleles	20	50.0	50.0	4.1	4.1
4 alleles segregating per line	10	-	39.8	-	4.1
	20	41.3	44.4	7.6	6.5
	50	-	14.6	-	12.6
2 alleles segregating per line	20	23.6	28.7	15.4	12.9

The principles of the multiple marker analysis of outbred line crosses can also be applied to the analysis of data from within a half-sib structured population (Knott *et al.*, 1994). In this application we consider there are a number of sires each with a group of progeny, which might be either daughters with records or sons with progeny test results (i.e. the 'daughter' or 'granddaughter' designs of Weller *et al.*, 1990). The probability of each progeny inheriting each of the sire gametes at a particular position in the linkage group can be calculated conditional on the genotypes of multiple markers. In fact for any progeny only the two nearest flanking markers which are informative need be considered (these may differ between progeny of the same sire depending upon the progeny genotypes and whether the dam

genotype is known). Within sires, phenotypic values can be regressed on the difference between the conditional probabilities and across sires the interaction of this regression with sire provides a test for the presence of a QTL very similar to that which would be used for a single marker (Weller *et al.*, 1990). Results of this approach and its comparison with the single marker approach show again that the use of multiple markers can increase power and provide better estimates of the QTL position and effect.

CONCLUSIONS

Interval mapping has proven to be a very useful tool for the detection and localisation of QTLs in populations derived from crosses between inbred lines. However, it is only now that problems, such as that of mapping multiple QTLs, have been recognised and methods are being developed to deal with them that the full power of interval mapping can be realised. As far as outbred lines are concerned, the power of interval mapping has yet to be fully harnessed. Because of the low information content of markers in outbred populations, using only the markers flanking an interval is not enough, and can lead to reduced power and biases in estimates of QTL effect and position. It is necessary to use all markers in a linkage group to fully utilise the information present. This approach could be implemented within a number of contexts - full ML, or mixed model (e.g. Goddard, 1992), or regression. The advantage of full ML is that it uses information on the distribution within marker classes which is neglected by the other methods. However, at least for inbred lines, omitting this information in a regression analysis does not result in very great loss of power. Regression approaches using multiple markers can in some circumstances be applied to outbred populations and can provide a useful increase in power and reasonable estimates of the effects and position of a QTL. It is likely that improvements in computer technology and methods such as Gibb's sampling will allow information from multiple markers in complex pedigrees to be fully utilised. However the methods are applied, future attempts to locate QTLs in livestock should attempt to harness the principles of interval mapping and utilise multiple markers in order to extract maximum information from the hard-won data.

ACKNOWLEDGEMENTS

We are grateful for support from the BBSRC, MAFF and the EC BRIDGE programme. Thanks to Peter for comments and to the boys for stimulating nights.

REFERENCES

ANDERSSON, L., HALEY, C.S., ELLEGREN, H., KNOTT, S.A., JOHANSSON, M., ANDERSSON, K., ANDERSSON-EKLUND, L., EDFORS-LILJA, I., FREDHOLM, M., HANSSON, I., HÅKANSSON, J. and LUNDSTROM, K. (1994) *Science* (in press).
DARVASI, A., WEINREB, A., MINKE, V., WELLER, J.I. and SOLLER, M. (1993) *Genetics* **134**: 943-951.
GODDARD, M.E. (1992) *Theor. Appl. Genet.* **83**: 878-886.
HALEY, C.S. and KNOTT, S.A. (1992) *Heredity* **69**: 315-324.
HALEY, C.S., KNOTT, S.A. and ELSEN, J.M. *Genetics* **136**: (in press)
HALEY, C.S. (1991) *Anim. Genet.* **22**: 259-277.
JANSEN, R.C. (1992) *Theor. Appl. Genet.* **85**: 252-260.
JANSEN, R.C. (1993) *Genetics* **135**: 205-211.
KNOTT, S.A. and HALEY, C.S. (1992a) *Genet. Res.* **60**: 139-151.
KNOTT, S.A. and HALEY, C.S. (1992b) *Genetics* **132**: 1211-1222.
KNOTT, S.A., ELSEN, J.M. and HALEY, C.S. (1994) These proceedings.
LANDER, E.S. and BOTSTEIN, D. (1989) *Genetics* **121**: 185-199.
LANDER, E.S. and BOTSTEIN, D. (1986) *Cold Spring Harbor Symp. Quant Biol.* **51**:49-62.
MARTINEZ, O. and CURNOW, R.N. (1992) *Theor. Appl. Genet.* **85**: 480-488.
NIEMANN-SORENSON, A. and ROBERTSON, A. (1961) *Acta Agric. Scand.* **11**: 163-196.
OTT, J. (1991) *Analysis of human genetic linkage*. Johns Hopkins University Press, Baltimore.
SIMPSON, S.P. (1992) *Theor. Appl. Genet.* **85**: 110-111.
VAN OOIJEN, J.W. (1992) *Theor. Appl. Genet.* **84**: 803-811.
WELLER, J.I., KASHI, Y. and SOLLER, M. (1990) *J. Dairy Sci.* **73**: 2525-2537.
WELLER, J.I. (1986) *Biometrics* **42**: 627-640.
ZENG, Z. (1993) *Proc. Natl. Acad. Sci. USA* **90**: 10972-10976.