Confidence Intervals in QTL Mapping by Bootstrapping

Peter M. Visscher, Robin Thompson and Chris S. Haley

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland

Manuscript received July 24, 1995
Accepted for publication February 24, 1996

ABSTRACT

The determination of empirical confidence intervals for the location of quantitative trait loci (QTLs) was investigated using simulation. Empirical confidence intervals were calculated using a bootstrap resampling method for a backcross population derived from inbred lines. Sample sizes were either 200 or 500 individuals, and the QTL explained 1, 5, or 10% of the phenotypic variance. The method worked well in that the proportion of empirical confidence intervals that contained the simulated QTL was close to expectation. In general, the confidence intervals were slightly conservatively biased. Correlations between the test statistic and the width of the confidence interval were strongly negative, so that the stronger the evidence for a QTL segregating, the smaller the empirical confidence interval for its location. The size of the average confidence interval depended heavily on the population size and the effect of the QTL. Marker spacing had only a small effect on the average empirical confidence interval. The LOD drop-off method to calculate empirical support intervals gave confidence intervals that generally were too small, in particular if confidence intervals were calculated only for samples above a certain significance threshold. The bootstrap method is easy to implement and is useful in the analysis of experimental data.

FOR many plant and animal species, genetic maps are available with a large number of highly polymorphic markers. Recently, statistical methods have been developed to use map information to detect quantitative trait loci (QTLs). These methods can be applied to data from crosses between inbred lines (LANDER and BOTSTEIN 1989; HALEY and KNOTT 1992), data from crosses between outbred lines (HALEY *et al.* 1994), and to outbred populations (FULKER and CARDON 1994; GEORGES *et al.* 1995). All these methods have been applied successfully to experimental and field data (*e.g.*, PATERSON *et al.* 1989; ANDERSSON *et al.* 1994; CARDON *et al.* 1994; GEORGES *et al.* 1995). There are still problems remaining with all these methods, for example, how to deal with multiple QTLs (JANSEN 1993, 1994; ZENG 1993, 1994), how to set a significance threshold when testing for a QTL (LANDER and BOTSTEIN 1989; CHURCHILL and DOERGE 1994; REBAI *et al.* 1994), and what the null hypothesis should be when testing for QTLs (VISSCHER and HALEY 1996).

All QTL mapping methods, whether based on maximum likelihood (LANDER and BOTSTEIN 1989) or on regression (HALEY and KNOTT 1992; MARTINEZ and CURNOW 1992) do not lend themselves to a straightforward calculation of a confidence interval (CI) for the QTL location. Such a CI is important in practice, because it may determine the strategies for further experiments to get closer to the QTL, or for using the QTL in breeding programs. For example, when using markers to introgress a QTL allele from a donor population into a recipient population, one strategy could be to introgress a donor marker haplotype that covers the 95% CI for the QTL. The methods used in practice (*i.e.*, the method used in widely used computer programs) to calculate CI for QTL locations is the so-called LOD drop-off method (LANDER and BOTSTEIN 1989). Using this method, the CI is calculated by finding the location at either side of the estimated QTL location that correspond to a decrease in the LOD score of 1 or 2 units. The total width corresponding to a 1 or 2 LOD drop-off is then taken as the confidence interval, and, asympotically, these should be approximately equivalent to 96.8 and 99.8% CI, respectively (MANGIN *et al.* 1994). However, using differences in likelihood between the estimated location of the QTL and locations elsewhere on the chromosome to determine approximate 90 or 95% CI may be biased for small and medium-sized populations (VAN OOIJEN 1992; MANGIN *et al.* 1994), because then the distribution of the test statistic does not truly follow a chi-square distribution. For example, VAN OOIJEN (1992) showed by simulation that the proportion of confidence intervals based on a support interval of 1 LOD that contained the QTL varied from 0.73 to 0.84, depending on the size of the QTL and the type (backcross or F_2) of population (and size). CI were calculated only for samples that gave significant evidence of a segregating QTL. Similarly, MANGIN *et al.* (1994) found by simulation that the 90% CI was biased downward, *i.e.*, the proportion of 90% CI that con-

Corresponding author: P. M. Visscher, Institute of Ecology and Resource Management, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JG, Scotland. E-mail: peter.visscher@ed.ac.uk

TABLE 1

Effect of number of bootstrap samples and heritabilities on confidence intervals

$h^{2\,a}$	n^b	CI90[c]	P90[d]	CI95	P95
0.01	50	89	0.98	94	0.99
	100	88	0.98	95	0.99
	200	87	0.98	94	0.99
	400	88	0.99	95	1.00
	800	89	0.99	95	1.00
0.05	50	62	0.93	73	0.96
	100	59	0.94	75	0.97
	200	60	0.95	72	0.98
	400	59	0.94	71	0.98
	800	59	0.94	71	0.98
0.10	50	41	0.93	52	0.96
	100	37	0.91	50	0.96
	200	39	0.91	49	0.96
	400	37	0.91	47	0.97
	800	39	0.90	49	0.95

Data for 200 backcross individuals were simulated with a 100-cM chromosome with markers every 20 cM. The QTL is located at 55 cM from one end of the chromosome.

[a] Heritability.

[b] Number of bootstrap samples.

[c] CIx is the mean width of the x% confidence interval (in cM).

[d] PIx is the proportion of the x% confidence intervals that contain the QTL.

tained the simulated QTL was <0.90, in particular for QTLs that explained a small amount of variation and for dense marker maps. For example, for a backcross population of 200 individuals, the empirical probabilities that the 90% CI based on the LOD drop-off method contained the actual location of the QTL was ~0.84 and 0.74 for a map density of 20 and 5 cM, respectively. Hence, the "one LOD drop off" concept as proposed by LANDER and BOTSTEIN (1989) is not recommended for use in practice, because the actual drop-off needed varies with each study and each QTL. Other methods to calculate the CI of QTL positions in populations derived from crosses between inbred lines have so far relied on simulation (VAN OOIJEN 1992; DARVASI *et al.* 1993; MANGIN *et al.* 1994). For the case of two flanking markers, MANGIN *et al.* (1994) derived complex analytical formulae for confidence intervals, assuming normality of residuals. Experimental QTL mapping populations are usually not large, typically in the range of 100 to 500 observations, and the distribution of the data may not be normal, nor will many assumptions made in simulation studies (*e.g.*, evenly spaced fully informative markers) hold in practice.

We propose use of a bootstrap method (EFRON 1979, 1982) to determine approximate confidence intervals for the position of QTLs in practice. The aim of this study is to test, by simulation, how well the bootstrap method works in QTL mapping experiments and to suggest how to use it in practice. Specific questions addressed were as follows. (1) How does the bootstrap CI compare to CI derived from the LOD drop-off method? (2) Does the proportion of CI that contain the QTL depend on the (true) position of the QTL on the chromosome? (3) What are the sizes of the 90 and 95% CI for relatively small experimental populations?

MATERIALS AND METHODS

Simulation: Data for N individuals (N = 200 or 500) from a backcross (BC) population derived from inbred lines were simulated. For each individual, single chromosomes of 100 cM with six or 11 (m = 6, 11) evenly spaced fully informative markers (including markers at the ends) were simulated. Hence, markers were spaced at either Δ = 10 or 20 cM. Crossovers were generated assuming HALDANE's mapping function without interference (HALDANE 1919). The chromosomes contained a single QTL, and its additive effect was determined so as to obtain heritabilities in the BC population of 1, 5, and 10%. Environmental residuals were normally distributed. The position of the QTL (d) was either at 15 or 55 cM from the start of the chromosome. For each set of parameters, 1000 replicate BC populations were simulated.

Model for analysis: Data were analyzed with a variation of the regression method of HALEY and KNOTT (1992). Their method fits a putative QTL at different places along the chromosome (*e.g.*, at 1-cM intervals) and calculates the test statistic at each point. The position giving the largest test statistic is the most likely position for a QTL. It can be shown (WHITTAKER *et al.* 1996) that instead of performing a search along the chromosome, identical answers can be obtained by performing a multiple regression of phenotypes on pairs of flanking markers and transforming the estimated effects of the two markers in each regression to estimates of the QTL effect and its location. So, instead of a search at 1-cM intervals, the search is over (m − 1) pairs of markers. This method was preferred because of its speed of calculation. Test statistics were calculated using a likelihood ratio test, assuming that residuals are normally distributed. Only a single putative QTL was postulated in the model fitted.

Bootstrapping: Bootstrap samples were created by sampling with replacement N individual observations. An observation consists of a marker genotype and a phenotype. So, at each bootstrap sample, we draw, with replacement, N observations out of the pool of (N) original observation. Some records can appear more than once in a bootstrap sample, while others are not included at all. After n bootstrap samples, the empirical central 90 and 95% CI of the QTL position was determined by ordering the n estimates and taking the bottom and top fifth and 2.5th percentile, respectively. Thus, for each simulated backcross population an empirical bootstrap confidence interval was obtained, and the simulated QTL lies either within or outside this empirical confidence interval. Averaged over replicate populations, the proportion of empirical bootstrap CI that contain the QTL was calculated. If the method works perfectly, this proportion would be 0.90 (0.95) when the 90% (95%) empirical bootstrap CI was determined. For each parameter set, 1000n (1000 replicates × n bootstrap samples) QTL mapping analyses were done.

LOD drop-off support interval: For limited combinations of parameters, confidence intervals were calculated also with the LOD drop-off method. In these cases, a putative QTL was fitted at 1-cM intervals, and a support interval was constructed by taking the interval corresponding to a drop in the test

TABLE 2

Effect of population size and heritability on confidence intervals derived from bootstrapping and the LOD drop-off method

		Bootstrap				LOD drop-off			
N^a	$h^{2\,b}$	CI90[c]	P90[d]	CI95	P95	CI90	P90	CI95	P95
200	0.01	87	0.98	94	0.99	68	0.83	80	0.92
	0.05	60	0.95	72	0.98	38	0.86	48	0.93
	0.10	39	0.91	49	0.96	23	0.83	29	0.92
500	0.01	76	0.96	86	0.99	56	0.87	68	0.94
	0.05	33	0.91	43	0.96	21	0.85	26	0.93
	0.10	18	0.90	24	0.95	13	0.86	16	0.94

Data were simulated with a 100-cM chromosome with markers every 20 cM. The QTL is located at 55 cM from one end of the chromosome. The number of bootstrap samples was 200.

[a] Population size.

[b] Heritability.

[c] CIx is the mean width of the x% confidence interval (in cM).

[d] PIx is the proportion of the x% confidence intervals that contain the QTL.

statistic of 2.71 and 3.84, for a 90% and 95% CI, respectively. As detailed in MANGIN *et al.* (1994), asymptotically these values correspond to a 90 and 95% support interval when the test statistic is distributed at a χ^2 with one degree of freedom. The values from the χ^2 correspond to a drop in the LOD of 0.58 and 0.83, respectively.

RESULTS

The number of bootstrap samples (n): For a QTL of large effect (giving a h^2 of 10%) at 55 cM for the start of the chromosome, results are presented in Table 1 for a population size of $N = 200$. It appears that in general there are only small differences in CI and proportions of CI that contain the QTL for different values of n, which is as expected, since the effect of n is on the variation of CI and not on their mean values. In some cases, results for small n (50) seem to be slightly different from those with $n > 50$. For the 90 and 95% CI there is little change for values for $n \geq 100$, erring on the side of caution, we used $n = 200$ in subsequent simulations. Nearly all empirical bootstrap CI were slightly conservatively biased, in that the proportion of 90 and 95% CI that contained the QTL usually were >0.90 and 0.95, respectively. Only for $h^2 = 0.10$ were the proportions very close to the desired values of 0.90 and 0.95.

The population size (N): In Table 2 two population sizes are compared for the bootstrap method and the LOD drop-off method. The CI are much smaller for the larger population, very roughly by a factor of 2. For the larger population size ($N = 500$), the empirical CI seemed to be unbiased for heritabilities of 0.05 and 0.10 for the bootstrap method. Suppport intervals for the LOD drop-off method generally were anti-conservative, *i.e.*, the probabilities that the 90 and 95% support intervals contained the QTL were <0.90 and 0.95.

Marker spacing (Δ): Using either a marker spacing of 10 or 20 cM for the QTL mapping did not result in very different empirical CI (Table 3).

Position of the QTL: Results for the position of the QTL at either 15 or 55 cM are presented in Table 4. It appears that the empirical CI from the bootstrap method is not sensitive to the position of the QTL.

Type-I error: For QTL with small effects the analysis will occasionally pick up spurious QTL elsewhere on the chromosome. These type-I errors are expected to bias the empirical CI. To investigate this, correlations

TABLE 3

Effect on marker spacing, population size, and heritabilities on confidence intervals

N^a	Δ^b	$h^{2\,c}$	CI90[d]	P90[e]	CI95	P95
200	10	0.01	85	0.99	93	1.00
		0.05	58	0.95	70	0.98
		0.10	37	0.92	47	0.96
	20	0.01	87	0.98	94	0.99
		0.05	60	0.95	72	0.98
		0.10	39	0.91	49	0.96
500	10	0.01	74	0.98	84	0.99
		0.05	33	0.92	43	0.97
		0.10	17	0.91	22	0.96
	20	0.01	76	0.96	86	0.99
		0.05	33	0.91	43	0.96
		0.10	18	0.90	24	0.95

Data are simulated with a 100-cM chromosome with markers every 10 or 20 cM. The QTL is located at 55 cM from one end of the chromosome. The number of bootstrap samples was 200.

[a] Population size.

[b] Marker spacing in cM.

[c] Heritability.

[d] CIx is the mean width of the x% confidence interval (in cM).

[e] PIx is the proportion of the x% confidence intervals that contain the QTL.

TABLE 4

Effect of QTL position, population size, and heritabilities on confidence intervals

N^a	d^b	$h^{2\,c}$	CI90[d]	P90[e]	CI95	P95
200	15	0.01	90	0.93	96	0.97
		0.05	66	0.90	77	0.95
		0.10	38	0.90	50	0.94
	55	0.01	87	0.98	94	0.99
		0.05	60	0.95	72	0.98
		0.10	39	0.91	49	0.96
500	15	0.01	81	0.91	89	0.96
		0.05	32	0.90	43	0.95
		0.10	16	0.90	20	0.94
	55	0.01	76	0.96	86	0.99
		0.05	33	0.91	43	0.96
		0.10	18	0.90	24	0.95

Data were simulated with a 100-cM chromosome with markers every 20 cM. The QTL is located at 15 or 55 cM from one end of the chromosome. The number of bootstrap samples was 200.

[a] Population size.

[b] Position of the QTL on the chromosome in cM.

[c] Heritability.

[d] CIx is the mean width of the x% confidence interval (in cM).

[e] PIx is the proportion of the x% confidence intervals that contain the QTL.

TABLE 5

Correlations between the test statistic, the absolute difference between the true and estimated QTL position, and the 95% bootstrap confidence interval

N^a	$h^{2\,b}$	δ^c	$r(T, \delta)^d$	$r(T, CI)^e$	$r(\delta, CI)^f$
200	0.01	25	−0.25	−0.74	0.22
	0.05	11	−0.22	−0.75	0.29
	0.10	7	−0.12	−0.59	0.25
500	0.01	17	−0.24	−0.76	0.23
	0.05	6	−0.13	−0.61	0.22
	0.10	4	−0.06	−0.45	0.18

Data were simulated with a 100-cM chromosome with markers every 20 cM. The QTL is located at 55 cM from one end of the chromosome. The number of bootstrap samples was 200.

[a] Population size.

[b] Heritability.

[c] Absolute difference between the true and estimated QTL position.

[d] Correlation between test statistic and the absolute difference between the true and estimated QTL position.

[e] Correlation between the test statistic and the width of the 95% confidence interval.

[f] Correlation between the absolute difference between the true and estimated QTL position and the width of the 95% confidence interval.

were calculated between the test statistic, the absolute difference between the true and estimated position of the QTL, and the width of the bootstrap CI. Results are shown in Table 5. The correlation between the absolute difference between true and estimated QTL location and the test statistic was in the range of −0.1 to −0.2, *i.e.*, the larger the test statistic, the smaller the deviation from the true QTL location. Correlations between the test statistic and the bootstrap 95% CI were larger, in the range of −0.5 to −0.8. For different population sizes and heritabilities, the correlation between the deviation from the true QTL location and the CI were very similar, with values of ~0.2.

In practice, CI will be calculated only for QTL effects that are deemed significant. Given the large negative correlation between the test statistic and the width of the CI, we would expect the average CI to be smaller and less biased for those experiments (or replicates) in which there is strong evidence for the presence of QTL. To investigate this, we calculated the CI for significant replicates only. First, the 95% percentile of the test statistic was calculated for the case of no segregating QTL ($h^2 = 0$), using 10,000 simulated populations. Thresholds that gave a type-I error of 5% for a marker spacing of 20 cM and for populations sizes of 200 and 500 were 6.8 for both population sizes (results not shown elsewhere), corresponding to a LOD of 1.5. Empirical CI were determined for the bootstrap method and the LOD drop-off method, only for those simulated populations in which the test statistic exceeded the threshold value (Table 6). Although the CI were still biased upward for the bootstrap method, the proportions of CI that contained the QTL were closer to the desired values of 0.90 and 0.95 than the CI based on all simulated populations (Tables 1–4). For the LOD drop-off method, empirical probabilities that the support intervals contained the QTL location were biased downward, in some cases severely so. For example, for heritabilities of 0.01, the proportion of 90% support intervals that contained the QTL were 0.57 and 0.80 for a sample size of 200 and 500, respectively, and 0.69 and 0.88 for the 95% support intervals (Table 6).

DISCUSSION

Empirical confidence intervals: A bootstrap method was used to determine empirical 90 and 95% confidence intervals for the location of a QTL. The method worked well, *i.e.*, the proportion of empirical CI that contained the QTL was close to expectation (either 90 or 95%), in particular if CI are determined only for those populations with significant QTL effects. The bootstrap method is extremely easy to implement and is very fast if used in conjunction with regression-based interval mapping. Performing 200 bootstrap samples for one simulated population took only about 2 CPU sec for $N = 200$ and ~5 CPU sec for $N = 500$ on a DEC Alpha-3000 workstation. This compares to hours of CPU time when using maximum likelihood (CHURCH-

TABLE 6

Effect of type-I errors on confidence intervals for the bootstrap and LOD drop-off method

			Bootstrap				LOD drop-off			
N^a	$h^{2\,b}$	Power[c]	CI90[d]	P90[e]	CI95	P95	CI90	P90	CI95	P95
200	0.01	0.21	61	0.94	75	0.97	30	0.57	40	0.69
	0.05	0.75	49	0.91	63	0.96	31	0.83	39	0.91
	0.10	0.98	36	0.90	46	0.95	23	0.83	28	0.91
500	0.01	0.40	59	0.94	73	0.97	33	0.80	43	0.88
	0.05	0.99	32	0.92	42	0.95	20	0.84	25	0.93
	0.10	1.00	18	0.90	24	0.95	13	0.86	16	0.94

Data were simulated with a 100-cM chromosome with markers every 20 cM. The QTL is located at 55 cM from one end of the chromosome. The number of bootstrap samples was 200. Confidence intervals were determined if the test statistic was above threshold 6.8.

[a] Population size.

[b] Heritability.

[c] Power is the proportion of simulations for which the test statistic was above 6.8.

[d] CIx is the mean width of the x% confidence interval (in cM).

[e] PIx is the proportion of the x% confidence intervals that contain the QTL.

ILL and DOERGE 1994). However, as these authors pointed out, the cost of running a computer for a couple of hours or days is trivial compared to the cost of collecting genotypes and phenotypes, and thus it would be reasonable to combine bootstrap resampling with the maximum likelihood analysis. We do not expect that using maximum likelihood instead of linear regression would alter our conclusions as the methods generally give very similar results (HALEY and KNOTT 1992; MARTINEZ and CURNOW 1992). We used regression because of its simplicity and speed of computation.

Although simulations were not performed using different assumption, *e.g.*, data from F_2 populations, unequal marker spacing, not fully informative markers, nonnormal data, there is no reason why the method cannot be applied to other kinds of data. Indeed, it is the flexibility of the bootstrap method that makes it attractive to use in practice.

Depending on how the information from a QTL mapping experiment is used subsequently, a confidence interval may be too crude a way to summarize the findings. For example, if the ultimate aim of an experiment is to clone the QTL, a precise location estimate is needed. In that case, a posterior distribution of the QTL location may be more appropriate than a simple confidence interval. An empirical posterior distribution of the QTL location can be obtained by bootstrapping. Figure 1 gives a standard representation of a test statistic plotted against the position along the chromosome for a single BC population of 200 individuals. The replicate was selected because of its shape, *i.e.*, near the maximum test statistic the curve appears bimodal (or trimodal). This was done to mimic situations in practice for which it is not clear whether there are two distinct peaks (QTLs) or whether the peaks are caused by a single QTL. Subsequently, 10,000 bootstrap samples were taken, and the frequency distribution of the QTL location was calculated. This empirical distribution is shown in Figure 2. It appears that the trimodal shape from Figure 1 is accentuated in the frequency distribution. There seems clear evidence that the QTL is not at the markers, while the probability of locating the QTL at either positions 50 or 70 cM is similar. The same conclusion could have been drawn from Figure 1, since the difference in the likelihood ratio test statistic is small between positions 50, 70, and 90. Other replicate populations showed empirical bootstrap distributions of different shapes, often with a clear single peak. One use of the empirical frequency distribution of the QTL location in practice might be to calculate the empirical probability that the QTL lies within each interval for use in marker-aided

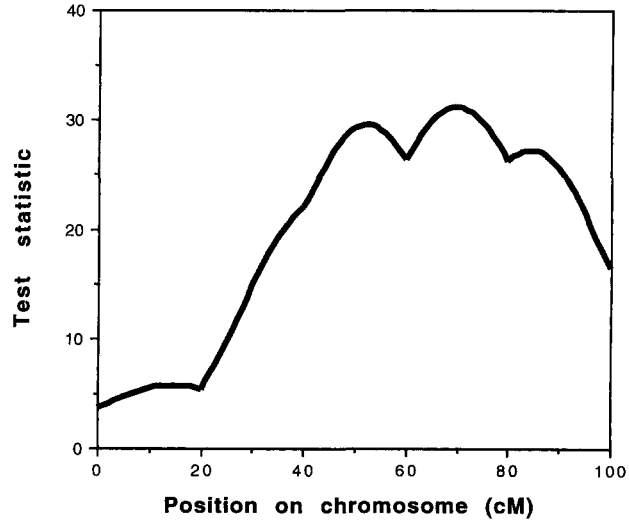

FIGURE 1.—Test statistic plotted along the chromosome for a single backcross population of 200 individuals. The QTL was at 55 cM and explained 10% of the phenotypic variance.

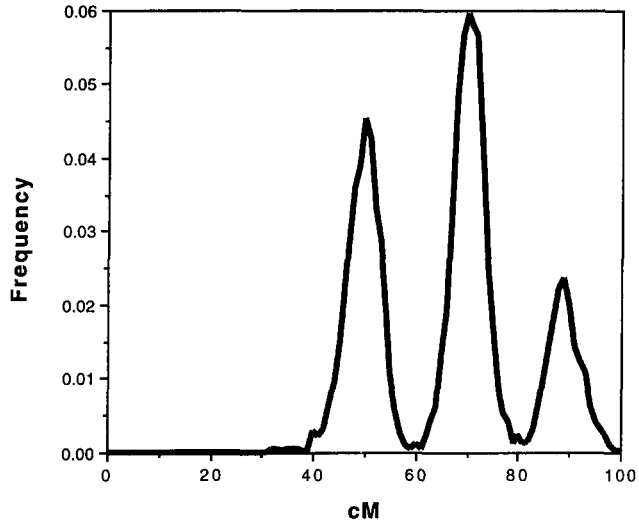

FIGURE 2.—Empirical frequency distribution of the estimated position of a QTL from 10,000 bootstrap samples from data of a single backcross population of 200 individuals (see Figure 1).

selection programs. For example, the empirical probabilities could be used as weights in marker-assisted selection breeding programs, thus taking account of the uncertainty about the location of a QTL.

The average width of the CI changed little with increasing marker spacing and decreased by more than the square root of 2.5 (500/200) when population size increased from 200 to 500. Under normality, we would expect the CI to be proportional to the square root of N. The actual dependence of the CI width of marker spacing and population size is possibly due to the change in the posterior distribution of the QTL location, which is nonlinear in Δ and N. Also, the CI has an upper bound of 100 cM, the length of the chromosome.

In the present study we investigated the use of the bootstrap technique only to estimate empirical CI for a single QTL. In practice, evidence for multiple QTLs per chromosome may be found either by employing a multidimensional search (HALEY and KNOTT 1992; MARTINEZ and CURNOW 1992) or by using the methods proposed by JANSEN (1993, 1994) and ZENG (1993, 1994). When using a direct search for two or more QTLs, application of the bootstrap seems obvious: for each bootstrap sample, one searches for the same number (q) of QTLs as detected in the original data sample, so that q locations are estimated in each bootstrap sample. After n bootstrap samples, either marginal or joint empirical CI can be determined using all location estimates. When using the method of JANSEN or ZENG, which fits cofactors in the model of analysis to condition on QTLs elsewhere in the genome, it becomes less clear from which pool to draw the bootstrap samples, because part of their method is to preselect which markers will be fitted in the model when performing the interval mapping. If the criterion to select markers as cofactors is kept constant across bootstrap samples, then by chance only one or even no intervals may be further explored to map the QTLs. Perhaps a logical way to apply the bootstrap in the cofactor method is to determine the empirical CI of one QTL at a time, by keeping the cofactors fixed across bootstrap samples. Hence, when detecting q QTLs, one would perform n bootstrap samples for each of the QTLs. However, further work is needed on finding CI for multiple QTL.

The aim of this study was to investigate the CI for QTL locations. However, the bootstrap method can equally be applied to other parameters in the model of analysis. To illustrate this, we calculated the empirical distribution of the additive QTL effect for $N = 200$, $h^2 = 0.10$ (allele substitution effect = 0.67), and $\Delta = 20$. The 90 and 95% CI widths for QTL effects were 0.50 and 0.62, respectively. The proportion of CI that contained the true QTL effect was 0.89 and 0.95, respectively. However, in these cases an empirical CI can be determined also by calculating the empirical standard error of the effect from the linear model (see, for example, DARVASI *et al.* 1993). Applying this procedure gave CI widths of 0.50 and 0.60, respectively, with 0.89 and 0.94 of empirical CI containing the true QTL effect. Hence, calculating the empirical standard errors directly from the data using the inverse of the $\mathbf{X}'\mathbf{X}$ matrix (regression) or information matrix (maximum likelihood) may be preferred. [Note that DARVASI *et al.* (1993) present the standard errors of the estimate of the mean of one the QTL genotypes in a backcross population in their Table 3. However, the expectation of that genotype class is the mean, so that the presented standard errors are just those pertaining to the overall mean, and not to the additive effect as suggested by the authors.]

Bootstrapping has been applied to other problems in genetics. For example, FELSENSTEIN (1985) has applied it to find confidence limits for phylogenies, and CHIANO and YATES (1994) have used bootstrapping in human genetic linkage analyses.

Bootstrapping *vs.* LOD drop-off method: Our results confirm those of VAN OOIJEN (1992) and MANGIN *et al.* (1994), in that the empirical CI obtained from a difference in the likelihood ratio generally is biased downward. Calculating support intervals only for samples that were deemed significant did not remove this bias for the LOD drop-off method, and in some cases made it worse (for $h^2 = 0.01$, compare Tables 2 and 6). In contrast, the bootstrap CI were generally less biased if only significant replicates were used to construct support intervals. We can explain this difference if we imagine a QTL with an infinitesimal effect, which would result in a proportion of significant samples approximately equal to the type-I error rate. Applying the LOD drop-off method for those, say, 5% of samples would

give a surface around the estimated QTL location that is too steep, giving support intervals that are too small. The (spurious) QTL are detected because of sampling variation, *i.e.,* by chance there is a difference in performance between individuals that have inherited markers from the two inbred lines. However, in the bootstrap method some of the chance associations are broken up, because at each bootstrap sample a different subset of the population is represented. Note that in this study a rather lax significance threshold was used because it was based on a single chromosome only. In practice, the significance threshold will be based on the entire genome, and the difference between the LOD drop-off method and the bootstrap method is likely to be larger.

Bootstrapping in linear models: The bootstrap method we choose, *i.e.,* sampling with replacement whole records (marker genotypes plus phenotypes), is only one out of many possible bootstrap strategies. There is no clear recipe on how to perform bootstrapping under linear models (WU 1986; HU and ZIDEK 1995). Different bootstrap algorithms may be used depending on whether (1) residuals in the linear model are independently and identically distributed, (2) variances are homogeneous or heterogeneous, or (3) effects that are fitted in the model are selected. Most studies so far have concentrated on (1) and (2) for a fixed linear model. When mapping QTL using linear regresssion, we are not strictly dealing with a fixed linear model with homogeneous residual variances, because the effects (markers) that are fitted in the model are a selected subset and because the actual QTL genotypes are not observed so that variances are heterogeneous. Only in the case of a very powerful experiment would the same markers (flanking the true QTL location) be selected that are fitted in the linear model. One argument against our sampling algorithm is that the design of the experiment varies over bootstrap samples (N.B. a different set of marker genotypes are contained in each bootstrap sample), and therefore the bootstrap CI of the QTL position includes variances due to different designs. An alternative sampling scheme is to calculate fitted values and residuals from the original data and then sample with replacement only residuals, so that the design is held constant over bootstrap samples (EFRON 1979; WU 1986). We tried this method and found that the CI were biased downward, *i.e.,* they were anti-conservative. This was particularly the case for QTL of large effects. For example, for $N = 500$ and $\Delta = 20$ and a QTL explaining 50% of the variance, the average 90 and 95% CI widths were 8 and 10 cM, respectively, and the proportion of CI that contained the QTL were 0.79 and 0.85, respectively. Thus it appears that this bootstrap sampling scheme did not create enough bootstrap variance of the QTL location. Until other bootstrap algorithms are found that perform markedly better, we prefer the algorithm proposed in this study because of its simplicity and ease of application.

Other resampling schemes: The permutation test suggested by CHURCHILL and DOERGE (1994) is another resampling scheme. However the purpose of the permutation test and the bootstrap method are different. The permutation test is used to determine an empirical threshold for significance testing, whereas the bootstrap method, as applied in this study, is used to determine an empirical confidence interval of a QTL location in the belief that there is an effect present. In practice, one could use both resampling methods: first, the permutation method to test for a significant QTL effect, and given significance, then the bootstrap method to find an empirical confidence interval for the location of that QTL. An alternative approach is to use bootstrappping for both significance testing and CI calculations, since given the empirical bootstrap distribution of the QTL effect a probability statement can be made analogous to a significance test (*e.g.,* the empirical probability that the QTL effect is zero). However, unlike the permutation test, this approach does not take multiple testing into account and would result in too many type-I errors.

P.M.V. was funded by the Marker Assisted Selection Consortium of the U.K. pig industry (Cotswold Pig Development Company Ltd., J.S.R. Farms Ltd., Newsham Hybrid Pigs Ltd., Pig Improvement Company/National Pig Development Company, and the Meat and Livestock Commission) and by Ministry of Agriculture, Fisheries and Food (MAFF), Department of Trade and Industry (DTI), and the Biotechnology and Biological Sciences Research Council (BBSRC). C.S.H. acknowledges support from MAFF, BBSRC and the European Commission. R.T. acknowledges support from MAFF. We thank BILL HILL and JOHN WHITTAKER for helpful comments.

LITERATURE CITED

ANDERSSON, L., C. S. HALEY, H. ELLEGREN, S. A. KNOTT, M. JOHANSSON *et al.,* 1994 Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science **263:** 1771–1774.

CARDON, L. R., S. D. SMITH, D. W. FULKER, W. J. KIMBERLING, B. F. PENNINGTON *et al.,* 1994 Quantitative trait locus for reading disability on chromosome 6. Science **266:** 276–279.

CHIANO, M. N., and J. R. W. YATES, 1994 Bootstrapping in human genetic linkage. Ann. Hum. Genet. **58:** 129–143.

CHURCHILL, G. A., and R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971.

DARVASI, A., A. WEINREB, V. MINKE, J. I. WELLER and M. SOLLER, 1993 Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics **134:** 943–951.

EFRON, B., 1979 Bootstrap methods: another look at the jacknife. Ann. Statist. **7:** 1–26.

EFRON, B., 1982 *The Jackknife, the Bootstrap and Other Resampling Plans.* Society for Industrial and Applied Mathematics, Philadelphia.

FELSENSTEIN, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791.

FULKER, D. W., and L. R. CARDON, 1994 A sib-pair approach to interval mapping of quantitative trait loci. Am. J. Hum. Genet. **54:** 1092–1103.

GEORGES, M., D. NIELSEN, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.,* 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907–920.

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

HALEY, C. S., and S. A. KNOTT, 1992 A simple regression method

for mapping quantitative trait loci in line crosses using flanking markers. Heredity **69:** 315–324.

HALEY, C. S., S. A. KNOTT and J.-M. ELSEN, 1994 Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics **136:** 1195–1207.

HU, F., and J. V. ZIDEK, 1995 A boostrap based on the estimating equations of the linear model. Biometrika **82:** 263–275.

JANSEN, R. C., 1993 Interval mapping of multiple quantitative trait loci. Genetics **135:** 205–211.

JANSEN, R. C., 1994 Controlling the type I and type II errors in mapping quantitative trait loci. Genetics **138:** 871–881.

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

MANGIN, B., B. GOFFINET and A. REBAI, 1994 Constructing confidence intervals for QTL location. Genetics **138:** 1301–1308.

MARTINEZ, O., and R. N. CURNOW, 1992 Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. **85:** 480–488.

PATERSON, A. H., E. S. LANDER, D. HEWITT, S. PETERSON, S. E. LINCOLN *et al.*, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

REBAI, A. R., B. GOFFINET and B. MANGIN, 1994 Approximate thresholds of interval mapping tests for QTL detection. Genetics **138:** 235–240.

VAN OOIJEN, J. W., 1992 Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. **84:** 803–811.

VISSCHER, P. M., and C. S. HALEY, 1996 Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. Theor. Appl. Genet (in press).

WHITTAKER, J. C., R. THOMPSON and P. M. VISSCHER, 1996 On the mapping of QTL by regression of phenotype on marker-type. Heredity (in press).

WU, C. F. J., 1986 Jackknife, bootstrap and other resampling methods in regression analysis. Ann. Stat. **14:** 1261–1295.

ZENG, Z-B, 1993 Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc. Natl. Acad. Sci. USA **90:** 10972–10976.

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

Communicating editor: Z.-B. ZENG