# Improved Confidence Intervals in Quantitative Trait Loci Mapping by Permutation Bootstrapping

**Jörn Bennewitz, Norbert Reinsch and Ernst Kalm**[1]

*Institut für Tierzucht und Tierhaltung, Christian-Albrechts-Universität, D-24098 Kiel, Germany*

Manuscript received July 30, 2001
Accepted for publication January 9, 2002

ABSTRACT

The nonparametric bootstrap approach is known to be suitable for calculating central confidence intervals for the locations of quantitative trait loci (QTL). However, the distribution of the bootstrap QTL position estimates along the chromosome is peaked at the positions of the markers and is not tailed equally. This results in conservativeness and large width of the confidence intervals. In this study three modified methods are proposed to calculate nonparametric bootstrap confidence intervals for QTL locations, which compute noncentral confidence intervals (*uncorrected method I*), correct for the impact of the markers (*weighted method I*), or both (*weighted method II*). Noncentral confidence intervals were computed with an analog of the highest posterior density method. The correction for the markers is based on the distribution of QTL estimates along the chromosome when the QTL is not linked with any marker, and it can be obtained with a permutation approach. In a simulation study the three methods were compared with the original bootstrap method. The results showed that it is useful, first, to compute noncentral confidence intervals and, second, to correct the bootstrap distribution of the QTL estimates for the impact of the markers. The weighted method II, combining these two properties, produced the shortest and less biased confidence intervals in a large number of simulated configurations.

SEVERAL statistical methods were proposed for detecting quantitative trait loci (QTL) using marker information. Most of them are based either on maximum-likelihood procedures (Lander and Botstein 1989) or on regression methods (Haley and Knott 1992; Martinez and Curnow 1992). However, none of these methods led to a straightforward calculation of a confidence interval. These are intervals on the chromosomes that contain the QTL with a given coverage probability $P$ (usually 90 or 95%). Calculating small confidence intervals that hold the stated coverage probability is of fundamental importance for practical breeding purposes as well as for molecular biologists. For example, breeders who wish to use the marker information in marker-assisted selection (MAS) breeding schemes need to know the optimum length of the chromosome segment of interest, and biologists need this knowledge for any further decisions regarding experimental fine-mapping strategy.

Different methods were described to calculate confidence intervals in QTL mapping. Conneally *et al.* (1985) and Lander and Botstein (1989) proposed the LOD drop-off method. This is based on the conversion of the distribution of the likelihood-ratio test statistic of linkage *vs.* no linkage between a marker and a putative QTL into a chi-square distribution. Using this method, the endpoints of the confidence interval are calculated by finding the locations at either side of the estimated QTL position having the maximum LOD score that have a LOD score of 1 unit less than the estimated QTL position. The confidence interval is then defined by all those points on the chromosome where the LOD score has fallen within 1 unit from the maximum (*i.e.*, the positions between the two interval endpoints). However, this method often produces intervals that are too short; *i.e.*, the intervals contain the QTL with a lower coverage probability than stated. Van Ooijen (1992) and Mangin *et al.* (1994) showed this to be the case for small and medium-sized populations and for QTL of small effects because the regular conditions for the conversion of the likelihood-ratio test toward a chi-square distribution do not hold in these genetic configurations. To attain a higher probability of containing the QTL, Van Ooijen (1992) suggested that the LOD drop-off confidence intervals should be based on two LOD differences. In a simulation study it was found that the sizes of these confidence intervals were large and variable (Van Ooijen 1992) and, in population sizes used in real QTL mapping procedures, unsatisfactory for QTL with smaller effects.

Mangin *et al.* (1994) and Mangin and Goffinet (1997) developed complex formulas for calculating confidence intervals. These are based on a maximum-likelihood-ratio test using statistics in which asymptotic distribution does not depend on the effect of the QTL. The major difficulty of this method is the computation of the correct threshold for the likelihood-ratio test for

[1]*Corresponding author:* Institut für Tierzucht und Tierhaltung, Hermann-Rodewald-Str. 6, D-24118 Kiel, Germany.
E-mail: ekalm@tierzucht.uni-kiel.de

which Mangin and Goffinet (1997) proposed an approximate analytical solution.

Visscher *et al.* (1996) introduced a nonparametric bootstrap resampling approach (Efron 1979; Efron and Tibshirani 1993) to construct a confidence interval for a QTL position. The evaluation of generated bootstrap samples provided a distribution of QTL position estimates along the chromosome that forms the base for the confidence interval calculation. Within a simulation study they compared the nonparametric bootstrap method with the LOD drop-off method. The nonparametric bootstrap procedure seemed a good alternative to the LOD drop-off for defining confidence intervals for QTL positions, but tended to be unsuitably large when the QTL effect was small.

To reduce the width of the estimated bootstrap confidence interval a selective nonparametric bootstrap method was proposed by Lebreton and Visscher (1998). Their selection was based on criteria related to the assumed genetic model. By comparing both methods in a simulation study they found that the selective bootstrap method produced smaller confidence intervals with a reduced conservativeness. Moreover, they showed that there is scope for improvement of the classical bootstrap approach for calculating confidence intervals as proposed by Visscher *et al.* (1996).

Walling *et al.* (1998) provided a parametric bootstrap resimulation method (Efron 1982) as an alternative to the nonparametric bootstrap approach and tested this method in a simulation study. As the authors themselves pointed out, the parametric bootstrap method produced poor results: The derived confidence intervals varied from conservative intervals when the QTL position was at the location of the marker to anticonservative intervals when the QTL position was between two markers.

A main result from the study of Walling *et al.* (1998) was the discovery of a bias in calculating confidence intervals from the distribution of QTL position estimates from nonparametric bootstrap samples. This bias is defined as a significant difference between the observed and the stated coverage probability and is deemed to have been caused by the locations of the markers. Eliminating this bias would probably result in smaller confidence intervals. As mentioned by Lebreton and Visscher (1998), an additional reduction of the interval width could be possible by computing noncentral confidence intervals. Taking this into consideration, we present three modified nonparametric bootstrap schemes to calculate confidence intervals: (i) noncentral, (ii) marker corrected, and (iii) noncentral and marker corrected. Within a simulation study these schemes were compared with the nonparametric bootstrap method presented by Visscher *et al.* (1996).

## MATERIALS AND METHODS

**Simulation:** A first generation backcross population consisting of one half-sib family and derived from two inbred lines was generated and contained 300 or 500 ($N$ = 300, 500) progeny. For each individual a single 100-cM chromosome ($L$ = 100 cM) evenly spaced with markers every 20 cM ($\Delta$ = 20 cM) was simulated. For calculating the crossover probabilities the Kosambi mapping function was used. Each chromosome contained a single segregating diallelic QTL with a position midway between two flanking markers ($d$ = 30, 50 cM), closer to one of the flanking markers ($d$ = 25, 55 cM), or directly at a marker position ($d$ = 20, 60 cM). The effect of the QTL was expressed in units of the polygenic standard deviation, and to cover a wide range of effects it was varied between 0.1, 0.5, and 1.0 ($a$ = 0.1, 0.5, 1.0). Additional simulations were carried out to investigate the impact of a longer chromosome ($L$ = 140 and 180 cM, $\Delta$ = 20 cM, $N$ = 300, 500 with $a$ = 0.5 and $d$ in the middle of the chromosome) and a higher marker density [$\Delta$ = 10 cM and 10/5 cM (additional markers at 35, 45, 55, and 65 cM), $L$ = 100 cM, $a$ = 0.5 with $N$ = 300, 500, respectively]. To avoid any confounding effects due to different detection power (Darvasi *et al.* 1993) the QTL position was located exactly midway between two flanking markers when marker spacing was reduced ($d$ = 55 cM when $\Delta$ = 10 cM and $d$ = 52.5 cM when $\Delta$ = 10/5 cM).

The total genetic variance was the sum of the polygenic variance and the variance due to the QTL. The polygenic genetic component was normally distributed. The heritability of the trait, defined as the polygenic variance divided by the sum of polygenic variance and the nongenetic variance, was assumed to be 0.25. Random nongenetic factors were normally distributed. For each simulated genetic configuration 1000 replicates were generated. The simulation program was written in FORTRAN 77, supplemented with routines from the NAG library (Numerical Algorithms Group 1990). QTL analyses were done by using the programs BIGMAP and ADRQLT (briefly described in Reinsch 1999) and were based on the multimarker regression described by Knott *et al.* (1994). Confidence intervals were calculated with FORTRAN 77 programs and SAS procedures (SAS Institute 1992).

**Generating the bootstrap samples:** A total of 250 nonparametric bootstrap samples for each genetic configuration were generated as described by Visscher *et al.* (1996). Briefly, for each single bootstrap sample $N$ observations out of the pool of $N$ original observations were drawn with replacement, so that each observation consisted of an individual's marker genotype and the corresponding phenotype. Some observations may appear more than one time in a bootstrap sample while others may be not included at all. Each single bootstrap sample was analyzed and the best estimate of the position of the QTL was recorded. After 250 bootstrap samples the distribution of the 250 estimates of the QTL position was derived by ordering them along the chromosome. Because the QTL is linked to two flanking markers, this distribution is termed the *linkage distribution*.

**Modeling marker impact on bootstrap distributions:** As reported by Walling *et al.* (1998) the reasons for the marker impact on the distribution of the bootstrap estimates for the QTL position are an accumulation of the type I error rate at the marker loci and the general trend of the mapping procedure to put the estimate of a QTL position toward a marker locus. In the following section a closer look at this marker impact is presented by a simple model, and the theoretical background for a marker correction is given. At first we focus on the expected distribution of the best estimates for a QTL position out of a bootstrap sample when a QTL may be present but not linked with any marker. This distribution is termed the *null distribution*, because every estimate indicates a type I error. In this null distribution the probability of each position achieving the best estimate would be

$$P(i)_{\mathrm{nd}} \propto P_{\mathrm{e}}^{*} S(i), \qquad (1)$$

where $P(i)_{nd}$ is the probability of obtaining the best estimate for position $i$ on the chromosome in the null distribution (nd), $P_e$ is the probability for each position achieving the best estimate assuming equal probabilities for each position, and $S(i)$ is a "selection coefficient" of position $i$ in the null distribution. Because $P_e$ is a constant with the same value for every position (*i.e.*, one divided through the number of possible QTL positions), it can be stated

$$P(i)_{nd} \propto S(i). \quad (2)$$

The selection coefficient describes the probability of each position to obtain the best estimate as a deviation from the equal distribution and is determined by the position and the informativeness of the marker loci. It would be highest at the marker loci and lower at the loci between the markers. In this model, it is assumed that the selection coefficient $S(i)$ is the same for both the linkage distribution and the null distribution. The probability of each position in the linkage distribution $[P(i)_{ld}]$ to obtain the best estimate would be

$$P(i)_{ld} \propto P(i)_{QTL} * S(i), \quad (3)$$

where $P(i)_{QTL}$ is the probability of each position $i$ achieving the best estimate when a QTL is linked with a marker and is, of course, influenced by the effect of the QTL but not by the markers *per se*. Following Equation 3, an approximate correction for the impact of the markers can be achieved by weighting the probability of each position from the linkage distribution with the reciprocal value of the corresponding selection coefficient:

$$P(i)_{QTL} \propto P(i)_{ld}/S(i). \quad (4)$$

The distribution resulting from the corrected probabilities $[P(i)_{QTL}]$ to attain the best estimate for the QTL location is termed the *corrected distribution*.

**Weighted method I:** In the weighted method the marker correction approach was followed. First, the probability in the linkage distribution to attain the best estimate $P(i)_{ld}$ and the selection coefficient $S(i)$ had to be determined for each position on the chromosome. The values for $P(i)_{ld}$ were the corresponding observed frequencies of best QTL estimates in the linkage distribution $[f(i)_{ld}]$:

$$P(i)_{ld} = f(i)_{ld}. \quad (5)$$

The linkage distribution was obtained as described above. The selection coefficients $S(i)$ were calculated by computing the null distribution for each simulated genetic configuration. This was achieved by setting the recombination rate between the simulated QTL and every marker locus to 0.5. The null distribution was then built by the best estimates out of all bootstrap samples and all replicates (250 estimates × 1000 replicates) for each simulated genetic configuration. Following Equation 2 the selection coefficient for each position on the chromosome was the corresponding frequency of the best estimates in the null distribution $[f(i)_{nd}]$:

$$S(i) = f(i)_{nd}. \quad (6)$$

For each replicate the distribution of the probabilities $P(i)_{QTL}$ (the corrected distribution) was calculated using (4).

The confidence intervals with a coverage probability $P$ of 95 and 90% were calculated by using the top and the bottom 2.5th and 5th percentiles as the upper and lower interval endpoints, respectively. The obtained confidence intervals were central.

**Weighted method II:** For the weighted method II the corrected distribution was computed in the same manner as for the weighted method I. Whereas for the weighted method I one-half of the given error rate of $1 - P$ was subtracted from each side of the corrected distribution during the computation of the confidence intervals, in the weighted method II an analog highest posterior density (HPD) method was used to obtain noncentral intervals when the corrected distribution is not tailed equally. This method was adapted from the standard HPD approach as described in Box and Tiao (1992) and is presented in detail in the appendix.

**Uncorrected method I:** This method was the original one as proposed by Visscher *et al.* (1996) and did not correct for the marker impact. The central confidence intervals were calculated directly from the linkage distribution by taking the top and bottom 2.5th and 5th percentiles of the linkage distribution as the upper and lower interval endpoints, respectively.

**Uncorrected method II:** In this method confidence intervals were computed from the linkage distribution with the analog HPD method described in the appendix. The calculated intervals were noncentral and not marker corrected.

The confidence intervals with the stated coverage probability $P$ of 95 and 90% were calculated for each replicate by the four methods described above, and it was assessed whether the QTL was located within these intervals or not. The rate of the confidence intervals for each simulated genetic configuration that did not include the simulated position of the QTL was termed the *noninclusion rate* (NI), and this should be not greater than $1 - P$. For each configuration the average width of the confidence interval out of 1000 replicates was also calculated, and this is ideally small.

## RESULTS

The initial results for all four bootstrap schemes are presented in Tables 1–4. Summarized over all genetic configurations analyzed, the uncorrected method I produced the largest and most conservative confidence intervals. A reduction of the interval width and conservativeness was possible using the uncorrected method II. Further reduction of the width was achieved using the weighted method I and, even more, by the weighted method II without lifting the noninclusion rate seriously above $1 - P$.

**QTL effect:** As expected, an increase of the QTL effect resulted in a smaller confidence interval for all bootstrap schemes (Table 1 and 2). It is very unlikely that a QTL with an effect of $a = 0.1$ (this QTL would explain only 0.13% of the total variance in our simulated populations) would be detected in real mapping experiments and, hence, the uncorrected method I produced very conservative confidence intervals, which cover almost the whole chromosome. For this low QTL effect each confidence interval produced by this method was >93 cM. However, the uncorrected method II (weighted method I, weighted method II) produced 90% intervals for this small QTL effect that were significantly <90 cM (85 cM, 80 cM) for the 90% confidence intervals and were less conservative. All confidence intervals for situations of $a = 0.1$ were only marginally influenced by the remaining genetic parameters.

Elevating the QTL effect from $a = 0.1$ to 0.5 (this QTL would explain 3.2% of the total variance) reduced

**TABLE 1**

**Effect of QTL position within marker interval, population size, and QTL effect on the confidence intervals for the different bootstrap methods—QTL position proximal**

| | | | $N^c$ = 300 | | | | $N$ = 500 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 95% interval | | 90% interval | | 95% interval | | 90% interval | |
| $d^a$ | $a^b$ | Method | Width[d] | NI[e] | Width | NI | Width | NI | Width | NI |
| 20 | 0.1 | Uncorrected I | 98 | 0.7 | 94 | 2.4 | 98 | 0.8 | 94 | 3.2 |
| | | Uncorrected II | 95 | 2.1 | 87 | 6.7 | 95 | 2.2 | 87 | 6.9 |
| | | Weighted I | 89 | 3.2 | 81 | 9.6 | 89 | 3.8 | 81 | 10.9 |
| | | Weighted II | 86 | 5.3 | 77 | 13.4 | 86 | 5.2 | 77 | 12.6 |
| | 0.5 | Uncorrected I | 86 | 0.8 | 77 | 2.1 | 74 | 1.9 | 62 | 4.2 |
| | | Uncorrected II | 82 | 1.5 | 70 | 5.0 | 70 | 3.3 | 58 | 5.6 |
| | | Weighted I | 77 | 4.0 | 67 | 8.5 | 66 | 5.3 | 54 | 9.9 |
| | | Weighted II | 72 | 4.0 | 60 | 8.9 | 60 | 4.7 | 48 | 8.0 |
| | 1.0 | Uncorrected I | 40 | 1.3 | 30 | 2.7 | 21 | 1.2 | 16 | 3.2 |
| | | Uncorrected II | 38 | 1.4 | 28 | 4.1 | 21 | 1.8 | 16 | 4.1 |
| | | Weighted I | 37 | 4.5 | 28 | 9.6 | 21 | 4.4 | 16 | 10.0 |
| | | Weighted II | 33 | 2.6 | 25 | 6.3 | 20 | 2.4 | 15 | 5.7 |
| 25 | 0.1 | Uncorrected I | 98 | 0.5 | 94 | 2.2 | 98 | 0.3 | 94 | 1.3 |
| | | Uncorrected II | 95 | 1.7 | 87 | 4.1 | 95 | 0.9 | 87 | 4.0 |
| | | Weighted I | 90 | 1.8 | 84 | 6.2 | 90 | 1.2 | 84 | 3.9 |
| | | Weighted II | 87 | 3.2 | 79 | 10.1 | 87 | 2.7 | 78 | 10.1 |
| | 0.5 | Uncorrected I | 87 | 1.6 | 78 | 3.7 | 78 | 3.0 | 67 | 5.3 |
| | | Uncorrected II | 83 | 3.0 | 72 | 6.9 | 74 | 4.2 | 62 | 8.5 |
| | | Weighted I | 77 | 2.1 | 66 | 4.9 | 69 | 3.4 | 57 | 6.0 |
| | | Weighted II | 73 | 3.3 | 61 | 9.1 | 64 | 4.0 | 51 | 8.2 |
| | 1.0 | Uncorrected I | 53 | 3.0 | 41 | 6.7 | 32 | 4.6 | 25 | 7.1 |
| | | Uncorrected II | 49 | 4.2 | 39 | 9.3 | 31 | 4.6 | 23 | 9.5 |
| | | Weighted I | 47 | 2.7 | 38 | 6.2 | 31 | 4.0 | 24 | 6.1 |
| | | Weighted II | 43 | 3.3 | 33 | 7.9 | 28 | 4.1 | 22 | 7.6 |
| 30 | 0.1 | Uncorrected I | 98 | 0.5 | 94 | 1.5 | 98 | 0.3 | 94 | 0.9 |
| | | Uncorrected II | 96 | 1.2 | 88 | 3.3 | 95 | 0.9 | 87 | 2.5 |
| | | Weighted I | 90 | 1.2 | 82 | 2.5 | 90 | 0.6 | 82 | 2.5 |
| | | Weighted II | 87 | 2.1 | 78 | 6.3 | 87 | 2.1 | 78 | 4.8 |
| | 0.5 | Uncorrected I | 88 | 1.3 | 79 | 2.8 | 80 | 2.4 | 69 | 4.5 |
| | | Uncorrected II | 84 | 2.1 | 73 | 6.3 | 76 | 3.2 | 64 | 6.4 |
| | | Weighted I | 77 | 1.9 | 67 | 5.6 | 69 | 2.8 | 58 | 5.1 |
| | | Weighted II | 73 | 3.6 | 62 | 8.1 | 65 | 3.6 | 53 | 7.3 |
| | 1.0 | Uncorrected I | 55 | 3.0 | 44 | 6.4 | 37 | 4.4 | 29 | 6.7 |
| | | Uncorrected II | 52 | 3.6 | 42 | 7.3 | 35 | 4.3 | 28 | 7.9 |
| | | Weighted I | 48 | 2.9 | 38 | 6.4 | 35 | 4.4 | 28 | 7.1 |
| | | Weighted II | 44 | 3.7 | 35 | 8.6 | 33 | 4.5 | 26 | 9.0 |

Data were simulated with a 100-cM chromosome and a marker spacing of 20 cM. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Position of the simulated QTL (in centimorgans from the start of the chromosome).

[b] QTL effect, expressed in units of the polygenic standard deviation.

[c] Population size.

[d] Average width (in centimorgans) of the confidence interval.

[e] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

the average 90% interval width by 15–25 cM. The respective change by stepping 0.5 to 1.0 (11.8% of the total variance) was 40–50 cM, depending on the population size and QTL position. This reduction was remarkably constant for the four methods.

**Marker spacing:** Changing marker distances from 20 to 10 cM (top of Table 2 and Table 3) and adding four additional markers around the QTL position (Table 3) resulted in a slightly reduced interval width for all bootstrap schemes. The reduction for the 90% confidence intervals varied between 2 and 4 cM. The corresponding changes were between 1 and 5 cM when marker distance was reduced to 10 cM and when additional markers were added around the QTL position, depending on the population size and the bootstrap scheme. No general dependency of the noninclusion rate on the marker density was observed for any method.

**Population size:** Tables 1–4 show that the elevation

**TABLE 2**

**Effect of QTL position within marker interval, population size, and QTL effect on the confidence intervals for the different bootstrap methods—QTL position central**

| | | | $N^c$ = 300 | | | | $N$ = 500 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 95% interval | | 90% interval | | 95% interval | | 90% interval | |
| $d^a$ | $a^b$ | Method | Width[d] | NI[e] | Width | NI | Width | NI | Width | NI |
| 50 | 0.1 | Uncorrected I | 98 | 0.0 | 95 | 0.2 | 98 | 0.0 | 95 | 0.6 |
| | | Uncorrected II | 95 | 0.1 | 87 | 1.1 | 95 | 0.5 | 87 | 1.4 |
| | | Weighted I | 90 | 0.1 | 82 | 1.3 | 90 | 0.1 | 82 | 0.7 |
| | | Weighted II | 87 | 1.3 | 78 | 2.6 | 87 | 0.4 | 77 | 2.0 |
| | 0.5 | Uncorrected I | 87 | 0.5 | 78 | 2.1 | 78 | 1.6 | 67 | 3.6 |
| | | Uncorrected II | 83 | 1.6 | 71 | 4.9 | 74 | 1.9 | 63 | 7.1 |
| | | Weighted I | 76 | 1.0 | 65 | 2.9 | 66 | 1.9 | 55 | 4.6 |
| | | Weighted II | 72 | 2.1 | 61 | 5.5 | 63 | 2.9 | 52 | 7.7 |
| | 1.0 | Uncorrected I | 53 | 4.6 | 43 | 7.3 | 36 | 2.9 | 28 | 6.9 |
| | | Uncorrected II | 52 | 5.0 | 41 | 9.0 | 34 | 4.0 | 27 | 7.1 |
| | | Weighted I | 47 | 4.6 | 38 | 7.3 | 34 | 2.6 | 27 | 6.5 |
| | | Weighted II | 44 | 5.1 | 35 | 8.3 | 32 | 4.2 | 25 | 8.2 |
| 55 | 0.1 | Uncorrected I | 98 | 0.1 | 94 | 0.4 | 98 | 0.2 | 94 | 0.7 |
| | | Uncorrected II | 95 | 0.3 | 87 | 1.9 | 95 | 0.6 | 86 | 2.5 |
| | | Weighted I | 90 | 0.2 | 82 | 1.5 | 90 | 0.4 | 82 | 0.7 |
| | | Weighted II | 86 | 1.1 | 76 | 2.9 | 86 | 0.6 | 78 | 3.3 |
| | 0.5 | Uncorrected I | 86 | 1.4 | 77 | 3.3 | 78 | 1.4 | 66 | 3.7 |
| | | Uncorrected II | 82 | 2.0 | 70 | 6.3 | 74 | 2.4 | 62 | 6.2 |
| | | Weighted I | 74 | 2.0 | 63 | 4.4 | 66 | 1.8 | 55 | 4.1 |
| | | Weighted II | 70 | 2.9 | 59 | 7.2 | 62 | 2.8 | 51 | 7.0 |
| | 1.0 | Uncorrected I | 49 | 3.7 | 38 | 7.7 | 33 | 3.7 | 25 | 8.1 |
| | | Uncorrected II | 47 | 4.9 | 36 | 10.1 | 31 | 5.3 | 24 | 9.2 |
| | | Weighted I | 43 | 3.2 | 34 | 7.0 | 31 | 3.0 | 24 | 6.8 |
| | | Weighted II | 40 | 4.2 | 31 | 9.4 | 29 | 3.9 | 22 | 7.8 |
| 60 | 0.1 | Uncorrected I | 98 | 0.3 | 94 | 1.0 | 98 | 0.3 | 94 | 1.0 |
| | | Uncorrected II | 95 | 0.9 | 87 | 1.9 | 95 | 0.5 | 86 | 2.1 |
| | | Weighted I | 89 | 0.7 | 81 | 2.4 | 89 | 0.5 | 80 | 1.7 |
| | | Weighted II | 86 | 1.8 | 77 | 4.1 | 85 | 1.1 | 76 | 4.1 |
| | 0.5 | Uncorrected I | 83 | 0.8 | 72 | 1.6 | 73 | 0.5 | 60 | 1.8 |
| | | Uncorrected II | 79 | 1.4 | 67 | 3.4 | 69 | 0.9 | 56 | 3.8 |
| | | Weighted I | 71 | 2.0 | 60 | 5.4 | 62 | 2.6 | 50 | 7.7 |
| | | Weighted II | 67 | 3.2 | 55 | 6.6 | 58 | 2.8 | 46 | 7.3 |
| | 1.0 | Uncorrected I | 40 | 1.5 | 30 | 4.2 | 22 | 1.2 | 17 | 2.2 |
| | | Uncorrected II | 38 | 2.3 | 28 | 4.1 | 21 | 1.5 | 16 | 3.0 |
| | | Weighted I | 36 | 5.2 | 27 | 10.7 | 22 | 4.8 | 17 | 10.2 |
| | | Weighted II | 33 | 3.7 | 24 | 6.8 | 20 | 2.5 | 15 | 4.8 |

Data were simulated with a 100-cM chromosome and a marker spacing of 20 cM. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Position of the simulated QTL (in centimorgans from the start of the chromosome).

[b] QTL effect, expressed in units of the polygenic standard deviation.

[c] Population size.

[d] Average width (in centimorgans) of the confidence interval.

[e] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

of the population size from $N = 300$ to $N = 500$ led to significantly smaller confidence intervals for all four bootstrap schemes. This reduction was only marginal for $a = 0.1$ but for $a = 0.5$ it was between 12 and 14 cM and for $a = 1.0$ it was between 10 and 15 cM. The uncorrected method I was most sensitive for increased population size as the reduction of the interval width was greatest. This might be due to the fact that intervals calculated with this method were largest and consequently there was the most space for reduction. It was not possible to assign a change of noninclusion rates to the effect of an increased population size for the uncorrected method II and for the weighted methods I and II; it appears that the intervals computed with the uncorrected method I were in general less conservative when the population size was increased.

In general, a gain of information, provided by a greater population size, resulted in smaller intervals. This is

**TABLE 3**

**Effect of an increased marker density on the confidence intervals for the different bootstrap methods**

| | | | $N^c$ = 300 | | | | $N$ = 500 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 95% interval | | 90% interval | | 95% interval | | 90% interval | |
| $\Delta^a$ | $d^b$ | Method | Width[d] | NI[e] | Width | NI | Width | NI | Width | NI |
| 10 | 55 | Uncorrected I | 84 | 0.5 | 74 | 2.7 | 76 | 2.1 | 64 | 3.6 |
| | | Uncorrected II | 79 | 1.6 | 68 | 5.5 | 71 | 3.0 | 59 | 7.7 |
| | | Weighted I | 73 | 1.5 | 61 | 4.1 | 64 | 2.0 | 53 | 5.9 |
| | | Weighted II | 68 | 3.0 | 57 | 6.1 | 61 | 3.8 | 49 | 8.1 |
| 10/5 | 52.5 | Uncorrected I | 82 | 1.1 | 71 | 3.2 | 70 | 0.4 | 59 | 2.3 |
| | | Uncorrected II | 77 | 1.8 | 64 | 5.6 | 67 | 1.5 | 54 | 5.2 |
| | | Weighted I | 72 | 1.2 | 61 | 3.8 | 62 | 1.2 | 50 | 4.2 |
| | | Weighted II | 68 | 2.6 | 56 | 6.8 | 58 | 2.7 | 47 | 6.6 |

Data were simulated with a 100-cM chromosome and a simulated QTL effect of $a = 0.5$ units of the polygenic standard deviation. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Marker spacing (10 = one marker every 10 cM, 10/5 = four additional markers every 5 cM around the simulated QTL position).

[b] Position of the simulated QTL (in centimorgans from the start of the chromosome).

[c] Population size.

[d] Average width (in centimorgans) of the confidence interval.

[e] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

consistent with other studies concerning QTL bootstrap confidence intervals (Visscher *et al.* 1996; Lebreton and Visscher 1998; Walling *et al.* 1998). The effect of the larger population size was stronger than the effect of the higher marker density.

**Length of chromosome:** The length of the chromosome had a large impact on the size of the confidence intervals and a small impact on the noninclusion rates for all bootstrap schemes; *i.e.*, the increase of the chromosome length resulted in larger intervals and, in most cases, in slightly lower noninclusion rates (Table 4). The absolute difference between the length of the intervals produced by the four methods increased slightly with a longer chromosome without changing the ranking order.

**QTL position:** The effect of the QTL position can be

**TABLE 4**

**Effect of the length of the chromosome on the confidence intervals for the different boostrap methods**

| | | | $N^c$ = 300 | | | | $N$ = 500 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 95% interval | | 90% interval | | 95% interval | | 90% interval | |
| $L^a$ | $d^b$ | Method | Width[d] | NI[e] | Width | NI | Width | NI | Width | NI |
| 140 | 70 | Uncorrected I | 120 | 0.6 | 105 | 1.5 | 107 | 0.8 | 90 | 2.3 |
| | | Uncorrected II | 114 | 1.3 | 99 | 4.6 | 102 | 1.5 | 84 | 5.4 |
| | | Weighted I | 104 | 0.6 | 88 | 2.7 | 91 | 1.5 | 74 | 3.9 |
| | | Weighted II | 99 | 1.9 | 82 | 5.6 | 86 | 2.4 | 69 | 6.2 |
| 180 | 90 | Uncorrected I | 151 | 0.5 | 132 | 1.6 | 137 | 1.1 | 117 | 2.8 |
| | | Uncorrected II | 144 | 0.5 | 124 | 3.0 | 130 | 1.7 | 110 | 3.6 |
| | | Weighted I | 133 | 0.6 | 112 | 2.1 | 119 | 1.5 | 98 | 2.6 |
| | | Weighted II | 126 | 0.9 | 105 | 4.1 | 113 | 1.6 | 92 | 3.0 |

Data were simulated with a marker spacing of 20 cM and a simulated QTL effect of $a = 0.5$ units of the polygenic standard deviation. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Chromosome length (in centimorgans).

[b] Position of the simulated QTL (in centimorgans from the start of the chromosome).

[c] Population size.

[d] Average width (in centimorgans) of the confidence interval.

[e] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

attributed to two factors: the QTL position within a marker interval (midway between two markers, closer to one marker, or at the same position as the marker) and the chromosomal QTL position, whether proximal or central. To avoid any confounding effects these two factors were analyzed separately.

It seems that when the QTL is located centrally (position 60 instead of 20, 55 instead of 25, and 50 instead of 30; Tables 1 and 2) the computed intervals are smaller, regardless of the method. This is in close agreement with Lebreton and Visscher (1998) and Walling *et al.* (1998) and may be caused by the fact that the mapping procedure tends to estimate the QTL position closer to the middle of the chromosome, as reported by Hyne *et al.* (1995). Additionally, the noninclusion rates were in general lower for $a = 0.1$ and $a = 0.5$ but higher for $a = 1.0$ for all methods.

Changing the QTL position within the marker interval from midway between two flanking markers toward a marker position (top, middle, and bottom of Tables 1 and 2) resulted in significantly smaller confidence intervals for all four methods. For example, the reduction for the 90% interval for $a = 0.5$ and $N = 500$ was between 1 and 2 cM when $d$ was changed from 50 to 55 cM and ~4 cM when $d$ was changed from 55 to 60 cM, regardless of the method. The uncorrected methods I and II produced intervals that were more conservative when the QTL and the marker positions are identical compared to situations with QTL positions midway between two flanking markers. This is consistent with the findings of Walling *et al.* (1998). The noninclusions were in general increased for the weighted method I when $d$ was moved toward a marker position whereas no general dependency could be observed between the QTL position and the noninclusion rates for the weighted method II.

Another remarkable topic might be the differences between the interval widths from the uncorrected methods I and II on the one hand and the differences between the weighted methods I and II on the other hand when moving the QTL closer to the start of the chromosome. To investigate the impact of the QTL position on these differences in detail, additional simulations with $N = 500$, $a = 0.5$, $\Delta = 20$ cM, $L = 100$ cM, and a QTL position at $d = 10$ or 15 cM were carried out. The average 90% interval widths computed with the uncorrected method II (weighted method II) were subtracted from the corresponding intervals computed with the uncorrected method I (weighted method I) for the genetic configurations described above and the same configurations but with $d = 20$ and 30 cM (Table 1) and $d = 50$ cM (Table 2). Additionally, the ratio of the subtracted upper tail and the subtracted lower tail of the linkage distribution and the corrected distribution during the computation of the 90% confidence interval with the analog HPD method were calculated. The results show that both the differences between the central and noncentral confidence intervals and the ratio of the subtracted upper and lower tail from the distributions increased when the QTL was located closer to the start of the chromosome (Table 5). The increase of the ratio was greater for the intervals computed from the corrected distribution than for those computed from the linkage distribution. When the QTL was midway on the chromosome ($d = 50$ cM) the ratio became, as expected, exactly one or near to one. It is important to note that the computed intervals for $d = 10$ cM were slightly anticonservative (10.4, 12.5, 12.9, and 13.9% noninclusions for the 90% intervals computed with uncorrected methods I and II and the weighted methods I and II, respectively; results not shown).

**TABLE 5**

**Effect of the QTL position on the differences between the widths of central and noncentral confidence intervals**

| | Differences in interval width[b] | | Ratio upper/lower tail[c] | |
|---|---|---|---|---|
| $d$[a] | UCM I − UCM II | WM I − WM II | ld | cd |
| 10 | 9 | 8 | 2.71 | 3.18 |
| 15 | 6 | 7 | 1.50 | 2.18 |
| 20 | 5 | 6 | 1.13 | 2.07 |
| 30 | 5 | 5 | 1.12 | 1.60 |
| 50 | 4 | 3 | 1.00 | 0.98 |

Data were simulated with a 100-cM chromosome, a marker spacing of 20 cM, and a simulated QTL effect of $a = 0.5$ units of the polygenic standard deviation. The population size was 500. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Position of the simulated QTL (in centimorgans from the start of the chromosome).

[b] Average differences of the 90% interval width computed by the uncorrected methods I and II (UCM I − UCM II) and by the weighted methods I and II (WM I − WM II).

[c] Ratio of subtracted upper tail and subtracted lower tail from linkage distribution (ld) and the corrected distribution (cd) during computation of the 90% confidence interval by the analog HPD method.

**Selection step:** In practical QTL mapping experiments confidence intervals would be computed only when the QTL is declared significant. To investigate whether the rejection of nonsignificant replicates would change the results a selection step for two genetic configurations similar to Visscher *et al.* (1996) was introduced. First, according to Churchill and Doerge (1994), threshold levels that gave a nominal type I error rate of 10 and 5% were calculated by taking the 90th and 95th quantiles of the test statistic of 10,000 analyzed simulated populations with no segregating QTL. Confidence intervals were computed only for replicates that showed QTL test statistics above the threshold. The results show that all computed confidence intervals were smaller and less conservative when introducing a selection step (top of Table 2 and Table 6) and when reducing the nominal type I error rate (Table 6). This finding is in agreement with Visscher *et*

**TABLE 6**

**Effect of rejection of nonsignificant replicates on the confidence intervals for the different bootstrap methods**

| | | | $N^b$ = 300 | | | | $N$ = 500 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 95% interval | | 90% interval | | | 95% interval | | 90% interval | |
| $\alpha^a$ | Method | Power | Width[c] | NI[d] | Width | NI | Power | Width | NI | Width | NI |
| 0.10 | Uncorrected I | 46.6 | 73 | 1.9 | 58 | 5.0 | 69.9 | 68 | 2.6 | 54 | 5.3 |
| | Uncorrected II | | 67 | 3.6 | 55 | 9.2 | | 65 | 3.4 | 52 | 8.7 |
| | Weighted I | | 61 | 2.2 | 48 | 6.5 | | 57 | 3.0 | 46 | 6.7 |
| | Weighted II | | 56 | 4.1 | 45 | 10.2 | | 53 | 3.7 | 43 | 9.5 |
| 0.05 | Uncorrected I | 37.0 | 70 | 2.1 | 55 | 5.9 | 58.7 | 65 | 3.0 | 51 | 7.2 |
| | Uncorrected II | | 65 | 4.3 | 54 | 9.8 | | 62 | 4.3 | 49 | 9.9 |
| | Weighted I | | 59 | 2.7 | 45 | 7.0 | | 54 | 3.6 | 44 | 7.8 |
| | Weighted II | | 54 | 4.3 | 43 | 11.1 | | 51 | 4.4 | 40 | 10.7 |

Data were simulated with a marker spacing of 20 cM and a simulated QTL effect of $a = 0.5$ units of the polygenic standard deviation. The QTL position was at 50 cM. Confidence intervals were computed only when the population showed a significant QTL. Experimental power was defined as the percentage of replicates that show a test statistic above threshold value corresponding to the nominal type I error rate. The number of bootstrap samples was 250 and the number of replicates for each genetic configuration was 1000.

[a] Nominal type I error rate.

[b] Population size.

[c] Average width (in centimorgans) of the confidence interval.

[d] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

*al.* (1996). Generally, the change of the nominal error rate from 10 to 5% reduced the interval width between 1 and 3 cM and increased the noninclusion rate by ~1–2%. No differences in sensitivity for the selection step between the four bootstrap methods could be observed.

**Number of bootstrap samples:** In computing confidence intervals from the linkage distribution Visscher *et al.* (1996) showed that there is only a small impact on the number of bootstrap samples. To examine whether this holds also for the computation of intervals from the corrected distribution, the number of bootstrap samples was varied for two genetic configurations ($\Delta = 20$ cM, $a = 0.5$, $L = 100$, $d = 50$ cM, and $N = 300, 500$, respectively). It seems that there is a lower bound for the number of bootstrap samples to generate the corrected distribution for interval computation that is between 100 and 200 and dropping it below 100 results in smaller intervals with a higher noninclusion rate. Increasing the number of bootstrap samples above 200 did not change the results (Table 7). Following this, the number of bootstrap samples (250) used in this study is appropriate.

## DISCUSSION

The classical bootstrap method (in this study the *uncorrected method I*) for calculating confidence intervals for a QTL position as proposed by Visscher *et al.* (1996) seems a suitable and practical approach, and in some recent real QTL mapping projects this method was used to obtain confidence intervals for the QTL position (*e.g.*, De Koning *et al.* 1998; Zhang *et al.* 1998; Walling *et al.* 2000). However, this method produces intervals that tend to be rather large and conservative, and therefore we proposed three methods (*uncorrected method II and weighted methods I and II*) to improve this classical approach to calculate smaller and less conservative confidence intervals and compared these methods within a simulation study with the classical approach.

**Bias of confidence intervals:** The bias of confidence intervals for QTL position estimates ($\text{bias}_{ci}$) can be defined as

$$\text{bias}_{ci} = \text{NI} - (1 - P),$$

where NI is the noninclusion rate. The $\text{bias}_{ci}$ for the intervals computed with the uncorrected method I was in general greatest and was upward except when the QTL was at the start of the chromosome and significantly increased for this method when the QTL position was located at a marker position. This is in agreement with the findings of Visscher *et al.* (1996) and Walling *et al.* (1998). When comparing the uncorrected method I with the weighted method I and the uncorrected method II with the weighted method II, respectively, it follows that the marker correction of the linkage distribution leads in general to a reduced bias of the confidence intervals as the noninclusion rate becomes closer to $1 - P$. Note that this bias was for the weighted method I, and even more, for the weighted method II, slightly downward, but not substantial, when the QTL effect was small and the QTL was located at the start of the chromosome but upward for the remaining configurations. Together with the fact that the weighted methods I and II produced significantly smaller confidence inter-

**TABLE 7**

**Effect of number of bootstrap samples on the confidence intervals for the weighted methods I and II**

| | | $N^b = 300$ | | | | $N = 500$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 95% interval | | 90% interval | | 95% interval | | 90% interval | |
| $B^a$ | Method | Width[c] | NI[d] | Width | NI | Width | NI | Width | NI |
| 30 | Weighted I | 70 | 2.0 | 62 | 5.1 | 61 | 5.2 | 52 | 10.1 |
| | Weighted II | 66 | 3.5 | 58 | 7.2 | 56 | 7.4 | 48 | 12.7 |
| 50 | Weighted I | 73 | 1.7 | 63 | 3.4 | 63 | 3.4 | 54 | 7.2 |
| | Weighted II | 69 | 3.0 | 59 | 7.0 | 60 | 5.5 | 50 | 9.8 |
| 100 | Weighted I | 74 | 1.2 | 65 | 3.1 | 65 | 2.5 | 54 | 4.9 |
| | Weighted II | 71 | 2.3 | 61 | 5.3 | 62 | 3.5 | 50 | 7.8 |
| 200 | Weighted I | 76 | 1.1 | 65 | 3.0 | 66 | 2.2 | 53 | 4.6 |
| | Weighted II | 72 | 2.1 | 61 | 5.3 | 63 | 3.3 | 52 | 7.6 |
| 400 | Weighted I | 76 | 1.2 | 65 | 2.9 | 66 | 2.1 | 54 | 4.5 |
| | Weighted II | 72 | 2.2 | 61 | 5.2 | 63 | 3.3 | 52 | 7.8 |

Data were simulated with a marker spacing of 20 cM and a simulated QTL effect of $a = 0.5$ units of the polygenic standard deviation. The number of replicates for each genetic configuration was 1000.

[a] Number of bootstrap samples.

[b] Population size.

[c] Average width (in centimorgans) of the confidence interval.

[d] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

vals than the uncorrected methods I and II it follows that the described model of characterizing the marker impact (Equations 1–4) and the marker correction in the weighted methods I and II is appropriate to obtain smaller and less biased bootstrap confidence intervals for the location of QTL that appear to be significant. Note that the probability of making a type I error is a function of the chosen significance level, which is not affected by any of the presented methods to compute confidence intervals.

**Marker correction:** Although it may be possible to calculate the influence of the markers on the linkage distribution, in terms of the selection coefficient for each possible QTL position (each centimorgan in our study), in an analytical way, in this study it was done empirically by computing the null distribution for each genetic configuration, *i.e.*, the distribution of the best QTL position estimates when there is no cosegregation between the QTL and any marker, and, hence, every estimate indicates a type I error per definition. This was possible because in the model it is assumed that the selection coefficient is the same for the linkage distribution as for the null distribution. The selection coefficient for each position was the corresponding frequency of the best estimate in the null distribution (Equation 6). In Figure 1A a null distribution for population size of $N = 500$, marker space of 20 cM, chromosome length of $L = 100$ cM, and QTL effect of $a = 0.5$ is presented. When the QTL is not linked with any marker, the probability for each position on the chromosome to achieve the best estimate is significantly higher at the marker loci. Over 45% of the estimates were placed there (under an equal distribution it would be only 5.94%). With an increased marker density (10 cM) 64% of the estimates would be placed at the marker loci. Except for the marker density and chromosome length it seems that the null distribution is of little variability, since varying the QTL effect and the population size resulted in almost the same null distribution (not shown).

The extent of the influence of the markers on the linkage distribution depends on the level of the QTL effect as it is shown in (3). Figure 1B shows an average linkage distribution from all bootstrap samples and all replicates for $N = 500$, $L = 100$ cM, $\Delta = 20$ cM, $a = 0.5$, and for QTL position of $d = 30$ cM from the start of the chromosome. When the QTL effect was $a = 0.1$ ($a = 0.5$, $a = 1.0$), 53% (32%, 9%) of the best estimates were located at the marker loci. The same genetic configurations but with a marker density of 10 cM gave roughly 10% higher values for the hits per marker locus, regardless of the QTL effect (results not shown), and this confirmed the presence of the marker influence on the linkage distribution as first reported by Walling *et al.* (1998). The two higher peaks at 20 and 40 cM are a result of the high value of the product of $P_{QTL}$ and selection coefficients at these positions (Equation 3). Both are relatively high because the distance to the position with the simulated QTL is only 10 cM and a marker is located at these positions, respectively. It follows from this and from the empirical results of the simulation (Tables 1 and 2) that the smaller the QTL effect (which is unknown *a priori* in real mapping experiments) and the lower the population size the more advisable it is to correct for the marker impact expressed in selection coefficients.

The correction for the marker was done by dividing

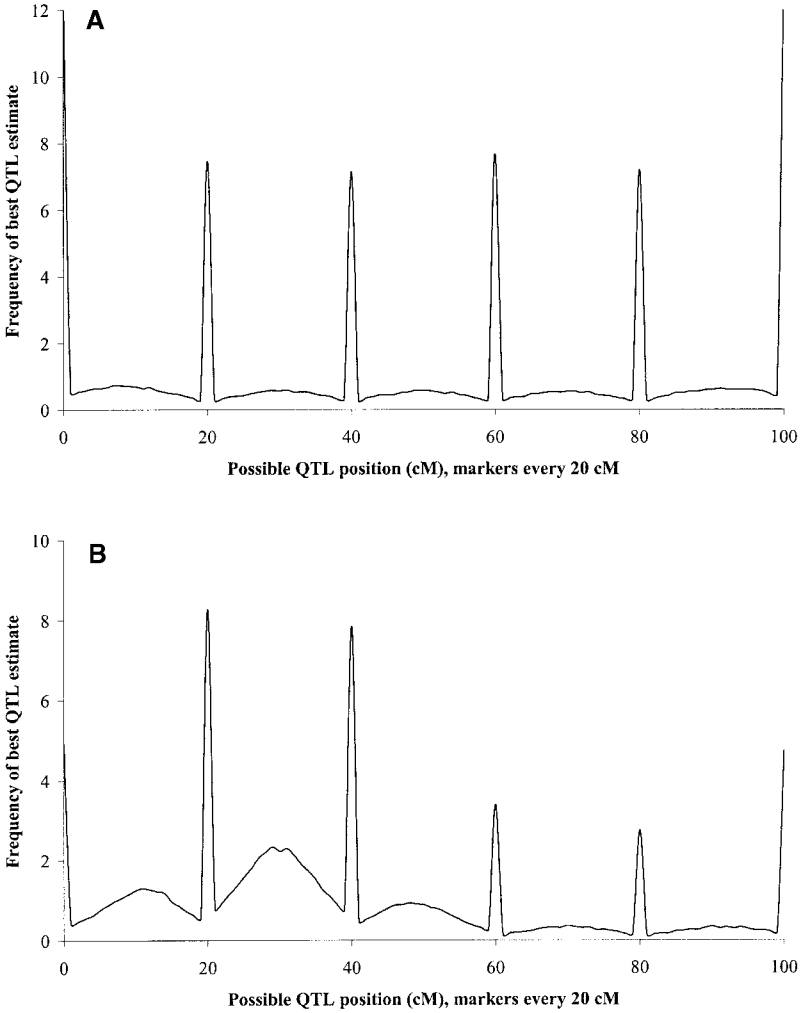
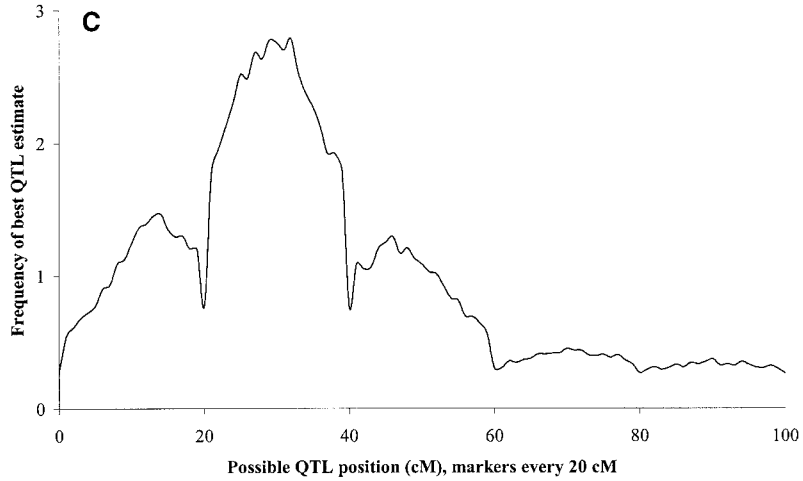


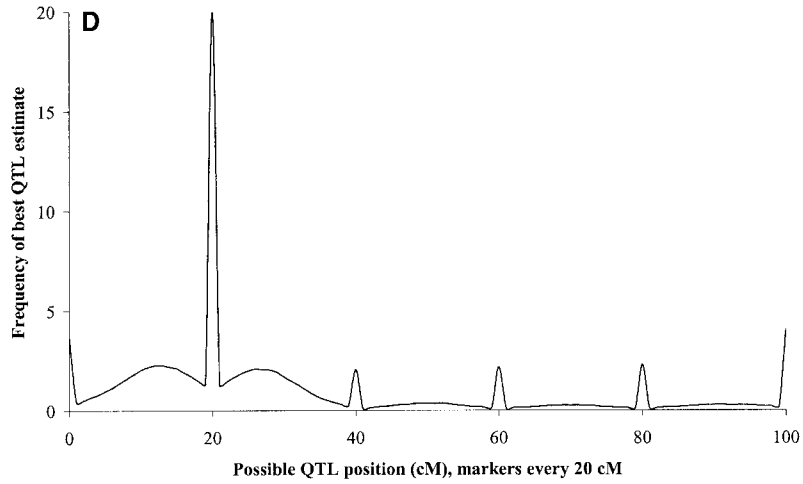

FIGURE 1.—Distributions of bootstrap QTL estimates along the simulated chromosome. (A) Empirical distribution of the QTL estimates when the recombination rate between the QTL and every marker was set to 0.5 (null distribution). (B) Empirical distribution of the QTL estimates when the QTL position was put at $d = 30$ cM (linkage distribution). (C) Empirical distribution of the marker-corrected QTL estimates when the QTL position was put at $d = 30$ cM (corrected distribution). (D) Empirical distribution of the QTL estimates when the QTL position was put at $d = 20$ cM (linkage distribution). (E) Empirical distribution of the marker-corrected QTL estimates when the QTL position was put at $d = 20$ cM (corrected distribution). A half-sib backcross family consisting of $N = 500$ individuals was simulated. The QTL effect was $a = 0.5$ and the markers were put at 0, 20, 40, 60, 80, and 100 cM, respectively. Data are from all bootstrap samples and all replicates ($250 \times 1000$ estimates).

the frequency of the best estimate in the linkage distribution by the value of the corresponding selection coefficient (see Equation 4) whereby the number of bootstrap samples should be >200 to compute a corrected distribution suitable for calculating confidence intervals (Table 7). Figure 1C shows an average corrected distribution again from all bootstrap samples and all replicates for $N = 500$, $\Delta = 20$ cM, $L = 100$ cM, $d = 30$ cM, and $a = 0.5$. The peaks at the marker loci observed in Figure 1, A and B, disappeared and the highest frequency of the best estimate was in the interval of the markers flanking the QTL. Right from position 60 cM the frequency became nearly equal for all positions (including marker positions).

An average linkage distribution for a situation QTL located at a marker position ($d = 20$ cM) is presented in Figure 1D ($N = 500$, $\Delta = 20$ cM, $L = 100$ cM, $a = 0.5$). The high peak at the QTL position can be attributed to the relatively high values of $P_{QTL}$ and to the selection coefficient at this point. In contrast to Figure 1B, the accumulation of the effect of a high $P_{QTL}$ value and a high selection coefficient is focused only on one position (20 cM) and this might be the reason for the shorter intervals with an increased conservativeness when computed with the uncorrected methods I and II. Correcting this average distribution results in Figure 1E. Right from position 60 cM the frequency of best estimates in this distribution became again nearly equal for all positions. Most of the probability mass is put between the markers at the start of the chromosome and position 40, but at the marker coinciding with the QTL position (20 cM) the frequency is significantly reduced (~1%). Similar patterns were observed at the positions of the markers flanking the QTL in the average corrected distribution in Figure 1C (QTL midway between flanking markers). Note that for this genetic configuration the null distribution follows the distribution presented in Figure 1A rather well and is therefore not shown.

**Computing the null distribution:** In the simulation the null distributions were derived in a simplified manner by setting the recombination rate between the QTL and any marker to 0.5. This had the advantage that the required CPU time was minimal, but is not possible in real QTL mapping experiments. Alternatively, the permutation approach as proposed by CHURCHILL and DOERGE (1994) can be used to compute the null distribution. The authors introduced this method to estimate empirical threshold values in QTL mapping, which are



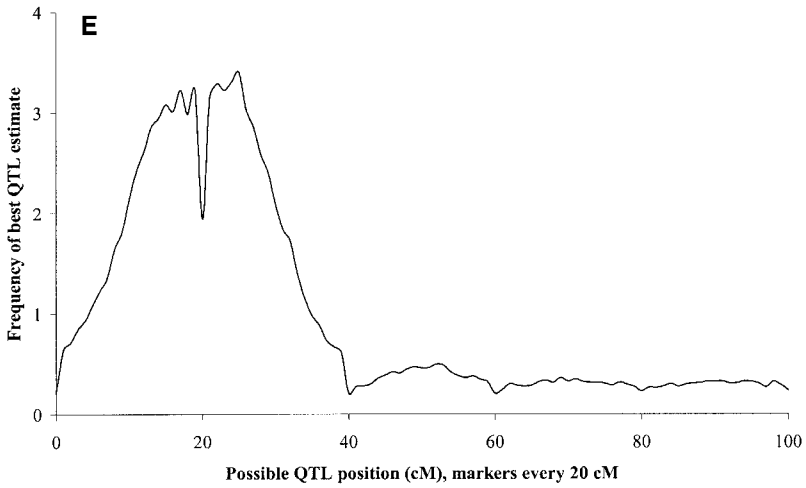




FIGURE 1.—*Continued.*

tailored to the experiment. By repeatedly randomly shuffling the trait values while keeping the marker data constant, they uncoupled the phenotype-genotype association, and after applying the QTL-mapping procedure, every estimate would indicate a type I error per definition. There is no reason why this distribution of the QTL estimates should not be influenced by the markers expressed in the selection coefficient for each position on the chromosome as described in our model, and so this distribution should be a very similar null distribution as obtained in this simulation. To show that the permutation method is suitable to calculate the null distribution we performed a further simulation with 1000 permutations of the phenotypic data for every replicate. The null distribution was the distribution from the QTL position estimates out of the evaluation of the permuted

**TABLE 8**

**Effect of the method for calculating the null distribution on the confidence intervals for the different corrected bootstrap methods**

| | | | 95% interval | | 90% interval | |
|---|---|---|---|---|---|---|
| nd[a] | $a$[b] | Method | Width[c] | NI[d] | Width | NI |
| Simp.[e] | 0.1 | Weighted I | 89 | 0.6 | 82 | 2.5 |
| | | Weighted II | 86 | 2.1 | 78 | 4.8 |
| | 0.5 | Weighted I | 69 | 2.8 | 58 | 5.1 |
| | | Weighted II | 65 | 3.6 | 53 | 7.3 |
| | 1.0 | Weighted I | 35 | 4.4 | 28 | 7.1 |
| | | Weighted II | 33 | 4.5 | 26 | 9.0 |
| Perm.[f] | 0.1 | Weighted I | 89 | 1.5 | 81 | 2.5 |
| | | Weighted II | 86 | 2.0 | 78 | 5.0 |
| | 0.5 | Weighted I | 68 | 2.5 | 58 | 7.0 |
| | | Weighted II | 65 | 4.5 | 53 | 10.0 |
| | 1.0 | Weighted I | 35 | 3.0 | 29 | 5.0 |
| | | Weighted II | 34 | 4.0 | 27 | 8.0 |

Data were simulated with a population size of $N = 500$ and with a marker spacing of 20 cM. The chromosome length was $L = 100$ cM and the QTL position was $d = 30$ cM from the start of the chromosome.

[a] Null distribution.

[b] QTL effect, expressed in units of the polygenic standard deviation.

[c] Average width (in centimorgans) of the confidence interval.

[d] Noninclusion rate, the rate of the confidence intervals that do not contain the real QTL position.

[e] Null distribution calculated and simplified by setting the recombination rate between the QTL and each marker to 0.5. The number of bootstrap samples was 250 and the number of replicates was 1000 for these genetic configurations.

[f] Null distribution calculated by permuting the phenotype data while keeping the genotype data constant. The number of permutations was 1000 for each replicate. The number of replicates for each genetic configuration was limited to 200.

data and was calculated for each replicate individually. To avoid a frequency of zero in the null distribution it was assumed that each position on the simulated chromosome would receive at least one QTL position estimate from the evaluation of the permuted data. The population size was $N = 500$, marker spacing was 20 cM, the QTL was put at $d = 30$ cM, and the QTL effects were $a = 0.1$, 0.5, and 1.0, respectively. To save CPU time the total number of replicates for each genetic configuration was limited to 200. The confidence intervals were calculated with the weighted methods I and II as described above and the obtained results were compared with the corresponding genetic configuration in the initial simulation, where the null distribution was derived in a simplified manner. The results are presented in Table 8 and show that there were no significant differences between the two simulated series although the null distribution was built out of 1000 estimates when computed with the permutation method, compared to 250,000 estimates in the initial simulation, and although different numbers of replicates were chosen. This confirmed the finding of the low variability of the null distribution as mentioned above.

In real QTL mapping experiments the permutation method has become a standard to calculate the error probabilities. Therefore, the null distribution can be derived as a by-product without extra computational work. It is only necessary to retain the position of the highest test statistic for each permutation and chromosome and to ensure that this distribution covers the whole chromosome. This is necessary to avoid a division by zero (Equation 4). This can be done by choosing either a sufficiently high number of permutations or, as in the present study, by adding one estimate to those positions not present in the initial distribution of the estimates from the evaluated permuted data. In the simulation, 1000 permutations were shown to be enough to compute marker-corrected confidence intervals for a 100-cM chromosome with six informative markers (Table 8) when adding additional hits. The number of permutations carried out to compute threshold values is in practice usually 10-fold higher.

**Analog HPD method:** The use of the analog HPD method to compute noncentral confidence intervals led to a reduced width and a reduced conservativeness of the intervals not only when the QTL was located closer to one end but also when located in the middle of the chromosome (Table 5). Intuitively one would not expect significant differences between central and noncentral intervals when the QTL is in the middle of the chromosome and the markers are located symmetrically on the chromosome. The likely explanation is that the bootstrap distributions, whether corrected for the markers or not, are tailed equally only in exceptional cases, regardless of the position of the QTL (not shown). A prerequisite for the analog HPD method to compute shorter confidence intervals is that the bootstrap distribution is not tailed equally (see APPENDIX and BOX and TIAO 1992).

**Conclusions:** In summary, it was shown that in the computation of bootstrap confidence intervals for estimated QTL locations it is useful, first, to correct the bootstrap linkage distribution for the marker impact and, second, to calculate noncentral intervals. Therefore, the use of the weighted method II is recommended, which combines the properties of the marker correction by the weighted method and the properties of the computation of noncentral intervals by the analog HPD method. The numerical results of the simulation showed that this method produced the shortest confidence intervals while maintaining approximately the noninclusion rate at $1 - P$ for a wide range of simulated genetic configurations. Provided permutation testing is used in an experiment, this method does not require extra computational work in comparison to the original bootstrap method. The combination of the permutation approach

and the bootstrap approach can be termed *permutation bootstrapping*.

This study has benefited enormously from the critical comments of two anonymous referees. It was supported by the German Cattle Breeders Federation (ADR) and the German Ministry of Education, Science, Research and Technology (BMBF).

## LITERATURE CITED

Box, G. E. P., and G. C. Tiao, 1992 *Bayesian Inference in Statistical Analysis*. Wiley Classics Library, New York.

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

Conneally, P. M., J. H. Edwards, K. K. Kidd, J. M. Lalouel, N. E. Morton *et al.*, 1985 Reports of the committee methods of linkage analysis and reporting. Cytogenet. Cell Genet. **40:** 356–359.

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.

De Koning, D. J., P. M. Visscher, S. A. Knott and C. S. Haley, 1998 A strategy for QTL detection in half-sib populations. Anim. Sci. **67:** 257–268.

Efron, B., 1979 Bootstrap methods: another look at the jackknife. Ann. Stat. **7:** 1–26.

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

Efron, B., and R. B. Tibshirani, 1993 *An Introduction to the Bootstrap*. Chapman & Hall, New York.

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.

Hyne, V., M. J. Kearsey, D. J. Pike and J. W. Snape, 1995 QTL analysis: unreliability and bias in estimation procedures. Mol. Breed. **1:** 273–282.

Knott, S. A., J. M. Elsen and C. S. Haley, 1994 Multiple marker mapping of quantitative trait loci in half-sib populations. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production. Vol. 21. Guelph, Canada, pp. 33–36.

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

Lebreton, C. M., and P. M. Visscher, 1998 Empirical nonparametric bootstrap strategies in quantitative trait loci mapping: conditioning on the genetic model. Genetics **148:** 525–536.

Mangin, B., and B. Goffinet, 1997 Comparison of several confidence intervals for QTL location. Heredity **78:** 345–353.

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.

Numerical Algorithms Group, 1990 *The NAG Fortran Library Manual*. Numerical Algorithms Group Ltd., Oxford.

Reinsch, N., 1999 A multiple-species, multiple-project database for genotypes at codominant loci. J. Anim. Breed. Genet. **116:** 425–435.

SAS Institute, 1992 *SAS Language*, Version 6, Ed. 1. SAS Institute, Cary, NC.

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

Visscher, P. M., R. Thompson and C. S. Haley, 1996 Confidence intervals in QTL mapping by bootstrapping. Genetics **143:** 1013–1020.

Walling, G. A., P. M. Visscher and C. S. Haley, 1998 A comparison of bootstrap methods to construct confidence intervals in QTL mapping. Genet. Res. **71:** 171–180.

Walling, G. A., P. M. Visscher, L. Andersson, M. F. Rothschild, L. Wang *et al.*, 2000 Combined analyses of data from quantitative trait loci mapping studies: chromosome 4 effects on porcine growth and fatness. Genetics **155:** 1369–1378.

Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen *et al.*, 1998 Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics **149:** 1959–1973.

Communicating editor: C. Haley

## APPENDIX: THE ANALOG HPD METHOD

Box and Tiao (1992) define a HPD region as a region in which the probability density of every point inside is at least as large as that of any point outside. The analog HPD method was adapted from the standard HPD to compute noncentral confidence intervals from the distribution of the best QTL estimate (from the linkage distribution in the uncorrected method II and from the corrected distribution in the weighted method II, respectively). When the distribution was not tailed equally these intervals were noncentral because unequal parts were subtracted from each tail. In contrast to the standard HPD method it does not make any assumptions regarding the form of the distribution, whether it is unimodal or even multimodal. If the distribution has the latter form the initial definition of a HPD region cited above does not hold.

The 95 and 90% confidence intervals were calculated iteratively as follows:

1. Draw a line parallel to the *x*-axis (the simulated chromosome) with an initial low value of ε (the distance between the line and the *x*-axis), label the position of the *x*-axis from the lower outstanding intercept point between the distribution and the line as $t_o$ and the corresponding position from the upper outstanding intercept point as $t_u$.
2. Calculate the sum of the frequencies of the distribution left from $t_o$ and right from $t_u$,

$$\text{sum} = \sum_{i=1}^{t_o} f(i) + \sum_{i=t_u}^{L} f(i),$$

where $t_o$ and $t_u$ are the positions of the *x*-axis from the lower and upper intercept points between the distribution and the parallel line, respectively, and $L$ is the length (in centimorgans) of the chromosome.

3. Decide three possible cases:

   Case 1. The sum is smaller than the given error rate of $1 - P$. In this case repeat steps 1 and 2 with a marginally increased value of ε, and label the new positions of $t_o$ and $t_u$ subsequently as $t'_o$ and $t'_u$, respectively.

   Case 2. The sum is equal to the given error rate of $1 - P$. The lower interval endpoint is the position of $t_o$ and the upper interval endpoint is the position of $t_u$. The frequencies of both interval endpoints are equal. See Figure A1A.

   Case 3. The sum is greater than the given error rate of $1 - P$ and too much was subtracted from the distribution. If it is the first round of

iteration start step 1 again with a lower initial value of ε. If it is not the first round of iteration this is not caused by a too low value for ε (a marginally lower value for ε would have resulted in case 2 and not in case 1 in the previous round of iteration); it is caused by the form of the distribution, which is in this case not unimodal. To determine the interval endpoints, two differences ($\Delta_1$ and $\Delta_2$) must be calculated:

$$\Delta_1 = t'_o - t_o$$
$$\Delta_2 = t_u - t'_u.$$

If $\Delta_1 < \Delta_2$ then the lower interval endpoint is the position of $t'_o$ and the upper endpoint (ue) is that point where

$$\sum_{i=1}^{t'_o} f(i) + \sum_{i=ue}^{L} f(i) = 1 - P.$$

This formula has to be solved iteratively. However, if $\Delta_2 < \Delta_1$, then the upper interval endpoint is the position of $t'_u$ and the lower endpoint (indicated by le) is that point where

$$\sum_{i=1}^{le} f(i) + \sum_{i=t'_u}^{L} f(i) = 1 - P. \qquad \text{(A1)}$$

Again, this formula must be solved iteratively. See also Figure A1B.

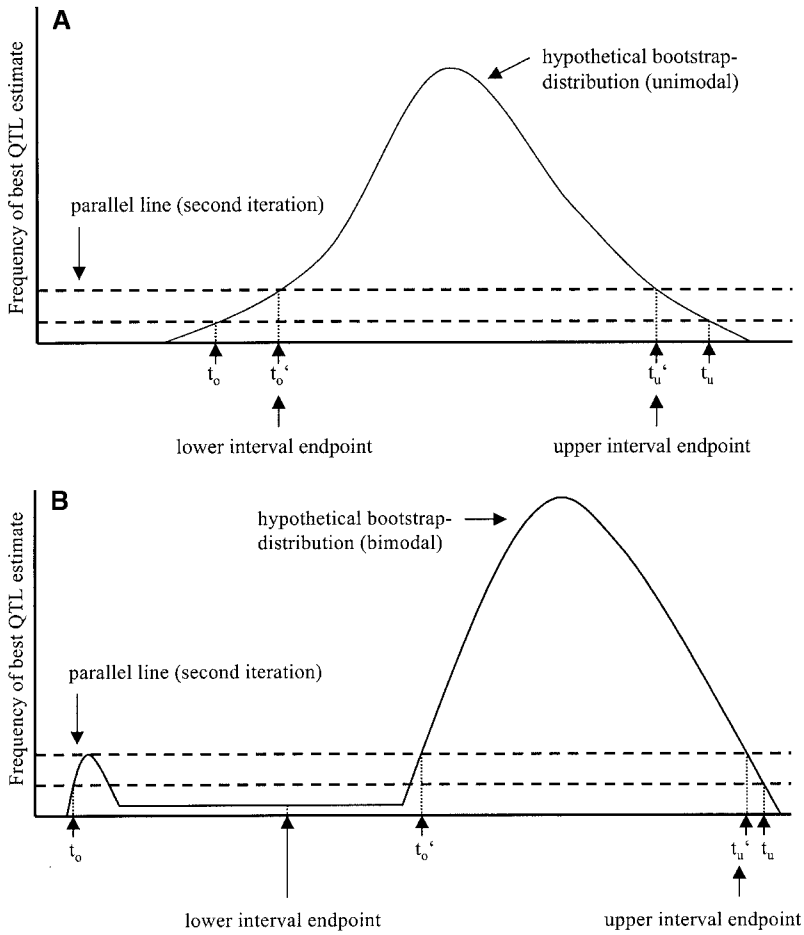



FIGURE A1.—Calculating confidence intervals from bootstrap distributions with the analog HPD method. (A) The hypothetical distribution is unimodal. The lower and upper interval endpoints are built by positions on the simulated chromosome of $t'_o$ and $t'_u$, respectively. (B) The hypothetical distribution is not unimodal. The upper interval endpoint is the position of $t'_u$ and the position of the lower interval endpoint is calculated iteratively using Equation A1.