A Nonparametric Bootstrap Method for Testing Close Linkage *vs.* Pleiotropy of Coincident Quantitative Trait Loci

Claude M. Lebreton,* Peter M. Visscher,† Christopher S. Haley,‡ Andrei Semikhodskii*,[1] and Steve A. Quarrie*

**John Innes Centre, Norwich NR4 7UH, United Kingdom, †Institute of Ecology and Resource Management, University of Edinburgh, Edinburgh EH9 3JG, United Kingdom and ‡Roslin Institute, Edinburgh EH25 9PS, United Kingdom*

Manuscript received December 30, 1997
Accepted for publication June 30, 1998

ABSTRACT

A novel method using the nonparametric bootstrap is proposed for testing whether a quantitative trait locus (QTL) at one chromosomal position could explain effects on two separate traits. If the single-QTL hypothesis is accepted, pleiotropy could explain the effect on two traits. If it is rejected, then the effects on two traits are due to linked QTLs. The method can be used in conjunction with several QTL mapping methods as long as they provide a straightforward estimate of the number of QTLs detectable from the data set. A selection step was introduced in the bootstrap procedure to reduce the conservativeness of the test of close linkage *vs.* pleiotropy, so that the erroneous rejection of the null hypothesis of pleiotropy only happens at a frequency equal to the nominal type I error risk specified by the user. The approach was assessed using computer simulations and proved to be relatively unbiased and robust over the range of genetic situations tested. An example of its application on a real data set from a saline stress experiment performed on a recombinant population of wheat (*Triticum aestivum* L.) doubled haploid lines is also provided.

THE gradual development of different types of molecular markers over the past decade has widened the applicability of quantitative trait locus (QTL) mapping to many species, even some for which variability between lines is limited. Increased marker density has facilitated accuracy of QTL positioning, which, in turn, allows the development of *comparative* QTL mapping. Comparisons of coincident associations of markers and QTLs can take place between genomes of related genus within a family, a subfamily, or a tribe, thus exploiting the conserved synteny between genomes that may exist. It can also take place between different traits and across different environments.

Among the different sorts of recombinant populations derived from line crosses, doubled haploid and recombinant inbred lines are more amenable to comparative QTL mapping between traits and between environments because the same genotypes can be used in several experiments. The experimental power is thus increased because the different trait distributions are no longer independent and can be studied as a multivariate *z*-dimensional distribution (where *z* is the number of traits or environments studied). Thus, as far as a QTL comparison between traits is concerned, existing genetic and environmental correlations can be exploited.

Corresponding author: Claude M. Lebreton, BIOCEM–Groupe Limagrain, 24, avenue des Landais, 63170 Aubière, France.
E-mail: claude.lebreton@bbsrc.ac.uk

[1] *Present address:* School of Biological Sciences, University of Manchester, Manchester M13 9PT, UK.

Ronin *et al.* (1995) and Korol *et al.* (1995) used the correlation between traits to increase the QTL detection power in experiments involving two traits with some correlation between them of genetic and/or environmental origin. In so doing, they assumed pleiotropy for every QTL mapped. In cases when this hypothesis was wrong, Jiang and Zeng (1995) demonstrated that the estimates of the QTL effects and positions were biased. While staying within the same framework, the latter authors extended the approach to the simultaneous analysis of more than two traits without having to assume pleiotropy for every QTL being mapped. They termed their method "Joint Mapping." They were then able to develop a test of close linkage *vs.* pleiotropy for a set of *n* QTLs, *i.e.*, one QTL per trait detected in the same genomic region for the *n* traits. For two traits, the test statistic is the likelihood ratio

$$\mathrm{LR}_2 = 2\ln(L_{20}/L_2), \qquad (1)$$

where L_{20} corresponds to the likelihood of the most likely full model, where the two QTLs are fitted without any constraints. L_2 is the likelihood of the model, where the two QTLs are fitted with the constraint that they must lie at the same position along the chromosome. Jiang and Zeng (1995) suggest that this likelihood ratio asymptotically follows a χ^2 distribution with 1 d.f. at the null hypothesis of pleiotropy. In the context of the LOD support interval of a QTL's position estimate, van Ooijen (1992) and Mangin *et al.* (1994) demonstrated that such a χ^2 distribution no longer constituted a good approximation of the real distribution of LR_2 when the

QTL was of small or medium effect. We are not aware that simulation studies have been performed to assess the distribution of LR_2 over a range of situations, so it is possible that the same applies to this test of close linkage *vs.* pleiotropy.

Visscher *et al.* (1996) proposed a less biased alternative to the LOD support interval. They estimated empirically a curve of probability density of the QTL's position using the nonparametric bootstrap, in which the original data are resampled N times with replacement and then plotted the frequency of the different positions observed. This procedure yielded confidence intervals on the QTL's position that performed better than LOD support intervals but tended to be too conservative when the QTL accounted for $<10\%$ of the total trait variance. Lebreton and Visscher (1998) managed to reduce the conservativeness and the size of the confidence intervals to a nearly unbiased level by conditioning the bootstrap on the genetic model, provided there was no bias on the QTL's position estimate.

In this study, we have used the selective bootstrap concept to develop a novel method to test close linkage *vs.* pleiotropy. In so doing, we replaced the likelihood ratio test statistic of Jiang and Zeng (1995) by a confidence interval on the estimated distance between two QTLs affecting different traits that are suspected to correspond to the same locus (hypothesis of pleiotropy). Our aim is to propose a test that is easy to implement and program (in the C-shell of a UNIX system, for example) with a variety of QTL mapping methods. The test must also be as little biased as possible over a range of genetic configurations commonly encountered. An absence of bias means that the null hypothesis is only rejected at the nominal frequency of type I error specified by the user. The bias and the power of the method are assessed by simulations using a range of genetic configurations. Finally, an example of its application to a real data set that comes from a saline stress experiment performed on a recombinant population of wheat (*Triticum aestivum* L.) doubled haploid lines is presented.

MATERIALS AND METHODS

QTL mapping procedures: The test of close linkage *vs.* pleiotropy that we propose can be implemented for any QTL mapping method that lends itself to an easy QTL model identification. However, as the method is based upon resampling, the speed of the QTL mapping procedure is important in allowing the assessment of the test on thousands of simulated data sets. Interval mapping-based methods, such as the interval mapping (*stricto sensu*) of Lander and Botstein (1989), the composite interval mapping of Zeng (1993, 1994) and Basten *et al.* (1996), or the multiple-QTL models of Jansen (1993) and Jansen and Stam (1994) provide asymptotically unbiased QTL parameter estimates due to the property of the maximum likelihood-based analysis. However, as these methods implement an iterative algorithm (the EM algorithm; Dempster *et al.* 1977) that must perform several cycles to converge to an estimate of the QTL parameters at each tested position of the putative QTL, they are slower than the multiple linear regression, for example, as simplified by Whittaker *et al.* (1996). Multiple linear regression constitutes an approximation of the maximum likelihood-based interval mapping in that only the within-marker class distribution of the residuals is considered and assumed normal (Martinez and Curnow 1992; Xu 1995). However, the method gives almost identical results to those from the maximum likelihood-based methods (Haley and Knott 1992). Walling *et al.* (1998) could nevertheless show, by using a large number of replicated simulations, the existence of a small but significant bias that tends to place the estimated QTL position closer to the nearest flanking marker. In this study we used two regression-based QTL mapping methods for the assessment of the test of linkage *vs.* pleiotropy.

The first QTL mapping procedure was used in situations of saturated coverage of the genetic map with one marker every centimorgan, to minimize any possible confounding effects due to the bias described above. Applying the principles established by Stam (1991), Zeng (1993), Rodolphe and Lefort (1993), and Wright and Mowers (1994), given the high and regular marker density over the genome, a simple marker-regressor selection was implemented to identify the QTLs. Because of their computational speed, regression methods lend themselves to resampling schemes. The QTLs were fitted at the positions of the selected markers, and therefore there were as many QTLs declared as there were markers selected. Their estimated additive effect was the partial regression coefficient related to the corresponding marker-regressor. During the marker selection, a first, relatively lax, forward procedure, with an inclusion F-ratio of 4.0, adds a subset of all the markers of the chromosome, some of which are included by chance (as discussed by Lebreton and Visscher 1998). Then, a stringent backward procedure was implemented to reject some markers. The stringency chosen determines the risk of type I error, *i.e.*, the risk of retaining a false QTL. The stringent F-threshold is calculated empirically according to the protocol presented in Lebreton and Visscher (1998) and follows the principles described by Churchill and Doerge (1994).

The second QTL mapping procedure was applied in situations where the robustness of the test was assessed with nonsaturated marker coverage of the genome. In a first stage, a subset of markers was chosen in the same way as the first method. Then, an "interpretation" stage followed, during which the number of QTLs and the sign of their additive effect (for the doubled haploid populations simulated in the context of this study) were estimated. A selected isolated marker, *i.e.*, a marker for which no adjacent markers were selected, was interpreted as evidence of a single QTL situated nearby. If two adjacent markers were selected with "apparent" effects (*i.e.*, the partial regression coefficient of the marker-regressor in the fitted multilinear model) of the same sign, a single QTL was assumed between the two markers. If the "apparent" effects of the two selected adjacent markers were of opposite signs, two QTLs were assumed nearby. In any case, the sign of the QTL effects was inferred to be that of the respective linked markers. This procedure constitutes a minimum estimate of the number of QTLs, but regardless, no linear method can fit more than one QTL per marker interval. This estimate of the number of QTLs on the chromosome of interest and the sign of their effect constitutes the basis on which the bootstrapped sample is rejected or not, as will be explained later. It is followed by a third stage, if the sample is retained, which consists of estimating the effect and position of the QTLs within marker intervals. Thus, if an isolated marker was selected in the first stage, both its flanking markers were tested in turn by including them in the model. The

marker that reduces the residual sum of squares the most is retained. The most likely interval that contains the QTL is thereby identified. The partial regression coefficients of the flanking marker-regressors fitted one at a time are used to estimate the QTL's additive effect and position within the interval according to Equations 2 and 3, as presented in Lebreton and Visscher (1998):

$$\boldsymbol{D_{j,k}} = \tfrac{1}{4}\ln \beta_{j+1} - \tfrac{1}{4}\ln \beta_j + \tfrac{1}{2}D_{j,j+1} \quad (2)$$

$$\boldsymbol{a} = \exp\left(\tfrac{1}{2}\ln \beta_j + \tfrac{1}{2}\ln \beta_{j+1} + D_{j,j+1}\right). \quad (3)$$

$\boldsymbol{D_{j,k}}$ is the estimated distance in centimorgans between the marker on the left and the QTL; $\boldsymbol{a}$ is the estimated additive effect of the QTL; β_j and β_{j+1} are partial regression coefficients of the trait phenotype on the marker-type, the markers j and $j + 1$ being fitted in turn; and $D_{j,j+1}$ is the distance in centimorgans between the flanking markers (j and $j + 1$).

In the case when two adjacent markers are selected, with apparent effects of the same sign, the most likely interval that contains the QTL is *de facto* identified. One of the flanking markers, say the one on the left of the interval, is removed, and the partial regression coefficient of the one on the right is calculated. Then the two markers are swapped over in the model, and the partial regression coefficient of the marker on the left is calculated. Then Equations 2 and 3 are applied to these coefficients. When the adjacent markers are of opposite signs the QTLs are fitted at the respective marker loci, and the apparent effects are retained as the estimate of the QTLs' effects—this as an *ad hoc* compromise between bias in QTL parameter estimation and speed of computation. If x ($x > 2$) adjacent markers were selected with "apparent" effects of the same sign, $x - 1$ QTLs in coupling would be assumed. However, our linear approximation was no longer able to estimate the QTLs' $2(x - 1)$ parameters (position and additive effect for each QTL) because there were only x values (the partial regression coefficients) to solve a system of $2(x - 1)$ equations with $2(x - 1)$ unknowns. Whittaker *et al.* (1996) also demonstrated the insolvability of this configuration when only the markers' "apparent" additive effects are available. Thus, compared to other regression-based QTL mapping methods that include other markers as background parameters so that the full model tested explains as much genetic variance as possible, such as MQTL (Tinker and Mather 1995), our method should yield very similar results, with only a slightly lower QTL detection power. This small decrease in QTL detection power is due to the fact that the pairs of adjacent markers are not systematically included in the compared models.

The test of close linkage *vs.* pleiotropy: The principle of the test consists of using bootstrap resampling to obtain an empirical estimate of the sampling distribution of the "distance" between the two QTLs—one per trait—that map in the same genomic region. Here we use the word *distance* in the loose sense because this quantity can actually take negative values and is defined as

$$d = \text{QTL position trait 2} - \text{QTL position trait 1}, \quad (4)$$

where QTL positions are measured in centimorgans from the leftmost marker on the chromosome. In the rest of this article, we shall use D as its estimate. For a chosen type I error X% for our test, *i.e.*, an X% risk of rejecting the null hypothesis of pleiotropy when pleiotropy was nevertheless the real configuration, a symmetrical X% confidence interval was calculated by taking the lower and upper $[(100 - X)/2]$th percentiles of the empirical distribution. If the interval did not contain the value "0," this value was considered as an outlier to the distribution of d, and the null hypothesis of pleiotropy was rejected. It accrues from the definition of d that it could vary within a range twice as large as the chromosome length, hence, in cases of poor experimental power, the span of a confidence interval on D could be larger than the chromosome length. Unlike in Jiang and Zeng (1995), QTL position trait 2 and QTL position trait 1 were estimated independently with our QTL mapping procedures. Note that, more generally, our nonparametric bootstrap test could be applied to the original data regardless of the QTL mapping procedure that is implemented.

In this article, for simplicity we considered only a population of doubled haploid lines. Thus, let $X_{n,m}$ be the matrix containing all the marker genotypes made of n row-vectors, n being the number of genotypes, and of m column-vectors, m being the total number of markers covering the genome. Also, let Y_1 and Y_2 be the vectors containing the n phenotypic values for trait one and trait two, respectively. The marker and the phenotypic data for the two traits of an individual are resampled jointly ($x_{i,1}, x_{i,2}, \ldots x_{i,m}, y_{1,i}, y_{2,i}$) and constitute one "resampleable" data point.

As in Lebreton and Visscher (1998), a selective nonparametric bootstrap was compared to a nonselective one. With the selective scheme, after the QTL mapping procedure was applied to the resampled data set, the parameter estimates contributed values to their estimated empirical distribution only if the QTL model, *i.e.*, the number of QTLs and the sign of their effects on the chromosome under study, conformed, for both traits, to those inferred from the original data set (see Figure 1B). Data sets producing other outcomes were rejected, and the original data were resampled again until N conforming outcomes were recorded. In the nonselective bootstraps, the same number of QTLs as that observed in the original data set was fitted whether they were significant or not in every resampled data set. All of the QTLs contributed values to the estimated empirical distributions of the parameters. Visscher *et al.* (1996) demonstrated that 150 bootstrap samples gave a sufficiently low bootstrap resampling error component (due to the limited number of resampled data sets as opposed to the error due to the limited size of the original data set itself; Efron and Tibshirani 1993). So we retained that value for the assessment of our test.

The general assessment protocol for our test: The aim of the protocol is to assess empirically the bias and the power of our test by evaluating the frequency of rejection of the null hypothesis of pleiotropy over many data sets that are the realizations of a given genetic configuration. The protocol was largely inspired by that described in Visscher *et al.* (1996) and is depicted in Figure 1A. In a real data set, it is trivial to say that the test of close linkage *vs.* pleiotropy would not be applied if either of the two QTLs tested did not show a significant effect. This is why, in our simulations, a prior selection was made in the original data sets (which we call *replicates* from now on) so that only the replicates that yielded the simulated number of QTLs and had effects of the right sign were kept. A genome-wise risk of 5% type I error in this prior selection was chosen to provide a stringency equal to what would realistically be applied in a real data set. Then the bootstrap test to be assessed was applied to each of these retained replicates with a specified risk of 5% type I error (risk of wrongly rejecting the null hypothesis of pleiotropy, as specified above) for all the simulations. The proportion of outcomes for which the null hypothesis was rejected constituted either: (1) an estimate of the power of the test when a situation of close linkage for the QTL pair studied was simulated, or (2) an estimate of the frequency of type I errors when pleiotropy was simulated, *i.e.*, an estimate of the bias of the method, because the unbiased expectation for this observed value should be the X% that was used to define the confidence interval on the value of d.

The genetic configurations tested: For simplicity, genomes

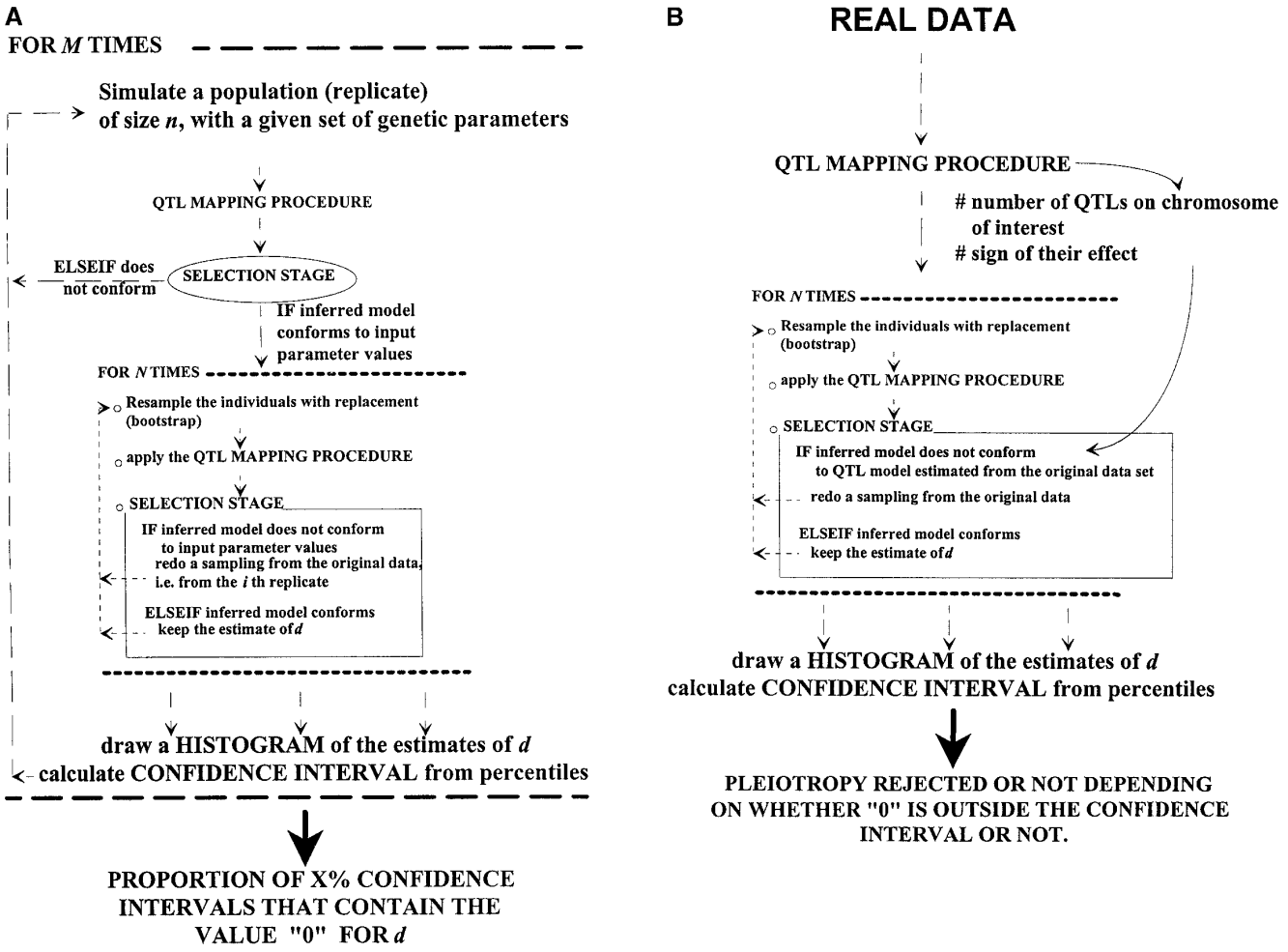

Figure 1.—Comparative diagram of the selective bootstrap test, assessed on simulated data (A) and applied on real data (B).

were made of one chromosome only, with a length of either 160 or 180 cM. Populations of doubled haploid lines of sizes ranging between 100 and 200 lines were simulated, with one QTL per trait in general. Heritabilities were defined as the ratio of the phenotypic variance of genetic origin between the line means over the total phenotypic variance between the line means. The trait's heritability was either 0.1 or 0.2. Only the mean of each line was simulated. Three QTLs per trait, on the same chromosome and in repulsion with a trait heritability of 0.6, were also simulated to assess the test in more complex genetic configurations. No epistatic interactions were simulated. The residuals were normally distributed.

Application of the test to real data: A recombinant population of 96 doubled haploid lines of bread wheat (*T. aestivum* L.) was used for some QTL studies on abiotic stress resistance by Quarrie *et al.* (1994). It is derived from the cross Chinese Spring × SQ1, with Chinese Spring being the maternal parent. SQ1 is a sibling line at F_7 of the line 25/3/2 from the cross Highbury × TW269 and was selected for high abscisic acid production under drought conditions. It is also an early line that flowers ~25 days before Chinese Spring in the absence of vernalization. Chinese Spring is known for its relative salinity tolerance.

Physiological, biochemical, and yield traits were measured on plants grown in a controlled environment and subjected to saline stress treatments comparing NaCl and Na_2SO_4, in two different experiments. In both experiments, five plants per genotype were used. The NaCl experiment was started at the end of May and finished in August 1994. The Na_2SO_4 experiment was started on 11 January 1996 and finished in the middle of April 1996. Before sowing, seeds were soaked in water, in Petri dishes at 4° overnight and then left at room temperature to germinate. When coleoptiles reached 1–2 mm, the seedlings were transferred to rockwool blocks. For the first 5 days, plants were grown in water that was then replaced with "Sangral" nutrient medium. The nutrient medium was gradually salinized when leaf one was fully expanded to reach a final concentration of 200 mm · liter^{-1} in NaCl and 100 mm · liter^{-1} in Na_2SO_4, respectively. Both experiments were performed in a greenhouse with a 16-hr day length. Daytime and nighttime temperatures were maintained at ~20° and 15°, respectively, and the air humidity was maintained at a level of 50–60%. Plants were harvested at maturity. The average spike length in the NaCl experiment and the length of leaf five in the Na_2SO_4 experiment were then measured. The data analysis was carried out on the means of each genotype.

The genome was covered with 331 markers, mostly RFLPs and AFLPs, plus a few morphological markers including the *Vrn1* gene (Galiba *et al.* 1995). This is the main gene controlling the vernalization requirements of European wheat varieties. It was scored as days to ear emergence in unvernalized plants. In our population, it produced a clear-cut segregation for the flowering time between the genotypes. Each of the 14 chromosomes of the A and B genomes were covered with 20

markers on average. However, due to a lack of variability in the D genome of *T. aestivum* L. as described by Cadalen *et al.* (1997), we could place only 3.7 markers per D genome chromosome.

The marker map was constructed using the software Mapmaker (version 3.0b) with the Haldane mapping function. A set of 21 nulli-tetrasomic lines, 1 for each chromosome, allowed us to assign unambiguously from one to eight markers to each chromosome as anchor points. Other markers were then grouped around these anchor points using a LOD threshold of 3.0 and a maximum recombination fraction of 0.3. Unlinked linkage blocks with anchor markers on the same chromosome were then forced into one linkage group and oriented relative to each other according to Gale *et al.*'s (1995) consensus map for the RFLP markers. Segregation distortion of the marker genotypes was investigated using a χ^2 test with 1 d.f. The normality of the trait distribution among the genotype means was checked using Shapiro and Wilk's *W*-statistic (Royston 1982). The QTLs were mapped, and the test of close linkage *vs.* pleiotropy was performed using the procedures described above that we implemented in a program in Fortran 90 (Ellis *et al.* 1994). No epistatic effects were fitted. The *F*-significance thresholds to retain markers in the backward selection stage of the QTL mapping procedure were determined empirically using 1000 permutations as described in Churchill and Doerge (1994) to achieve a target risk of 10% type I error genome-wise. Then, if the QTL was isolated (*i.e.*, no QTLs in the adjacent intervals), the *F*-statistic of a QTL was calculated by comparing the full model, including the markers flanking the QTL, to the reduced one that did not contain them. Otherwise, if the QTL was fitted at the marker position in the specific cases described above, only this marker was removed from the full model to calculate the *F*-statistic.

The output of our test of close linkage *vs.* pleiotropy was a *P* value calculated as twice the minimum percentage of outcomes with an estimated *d* either above or below zero. It corresponds to the integrated area of the empirical distribution's tail beyond zero and is multiplied by two because our test is two-tailed; it is the equality of a variable (*d*) to a specific value (zero) that is tested. Thus, *P* is the maximum risk of type I error one would need to adopt to be able to reject the null hypothesis of pleiotropy.

RESULTS

Biases in conservativeness of the selective bootstrap: Table 1 presents the results of a series of simulations aimed at assessing the bias of the selective bootstrap method to test close linkage *vs.* pleiotropy in a low power experimental design (although the heritability of an individual QTL can reach 20%). Nine genetic configurations were explored. In all of them, only one QTL was simulated per trait, for two traits, on a genome composed of a 180-cM-long chromosome. Because it was the bias that was being explored, the QTLs were simulated at the same position for both traits. The simulated values for the additive effects were equal to unity. The correlation between the residuals, *i.e.*, the environmental correlation between the traits, was equal to zero. Replicates were created until 2000 had been retained that conformed to the model simulated. For example, when a QTL of 10% heritability was simulated for both traits, 2260 replicates on average were generated to retain 1000 of them. The total number of replicates fell

TABLE 1

Frequencies of nonrejection of the null hypothesis of pleiotropy and sizes of the confidence intervals of the estimated distance between two QTLs for pleiotropic QTLs

					Select (20 cM)		Select saturated		Nonselect	
No.	Pos A	Pos B	H2A	H2B	*P* (95%)	CI size	*P* (95%)	CI size	*P* (95%)	CI size
1	90	90	10	10	97.7	102.84	95.7	113.7	99.2	146.0
2	15	15	10	10	96.4	114.83	96.1	121.2	98.8	195.8
3	5	5	10	10	94.0	108.7	95.8	126.9	98.5	210.9
4	5	5	10	20	88.4[a]	87.24	92.2[a]	98.2	96.6[a]	165.8
5	90	90	10	20	95.6	91.22	97.8	96.9	98.8	121.8
6	5	5	20	20	95.8	71.48	92.2	67.7	92.3	123.5
7	90	90	20	20	96.2	62.94	96.3	74.5	97.9	91.4
8	5	5	15	7.5	93.2[a]	110.3	94.6[a]	117.1	97.6[a]	198.3
9	90	90	15	7.5	94.4	103.8	96.0	110.4	99.2	139.7

No., number of the genetic configurations simulated; Pos A, position of QTL for trait A in centimorgans from the leftmost marker; Pos B, position of QTL for trait B in centimorgans from the leftmost marker; H2A, heritability of trait A; H2B, heritability of trait B; Select (20 cM), selective bootstrap with one marker every 20 cM; Select saturated, selective bootstrap with one marker every centimorgan; Nonselect, nonselective bootstrap with one marker every centimorgan; *P* (95%), frequency (percentage) of nonrejection of the null hypothesis (pleiotropy), using 95% confidence intervals on the estimated difference (QTL position trait B − QTL position trait A); CI size, average width of the 95% confidence intervals, in centimorgans. Both *P* (95%) and CI size were estimated over 2000 selected replicates. The simulated populations were made of 100 doubled-haploid lines. Five percent is the theoretical (expected) frequency of type 1 errors, and a significant departure of the *P* (95%) observed frequency from the nominal 95% indicates a bias of the test in the particular configuration.

[a] Configurations for which there is a significant bias in the estimate of the difference (QTL position trait B − QTL position trait A).

to 1200 when a QTL of 20% heritability was simulated for both traits (results not shown). A saturated marker coverage of one marker every centimorgan was compared to a density of only one marker every 20 cM.

The percentage estimates of inclusion of the value "0" by the 95% confidence intervals for the estimate of *d* were within confidence limits of ±1.3%. The estimates of the QTL additive effects were significantly biased upward due to the selection of the replicates and then of the bootstrapped samples. This bias is very small for QTL heritabilities of 20%: +5.6% ± 2% on average over the replicates and +9.5% ± 2% on average over the bootstrap samples, for both marker densities. For a heritability as low as 7.5%, the bias became more important: +33% ± 2% on average over the replicates and +49% ± 2% on average over the bootstrap samples with the dense marker coverage (results not shown). Biases were significantly higher with the sparse coverage of one marker every 20 cM: +40% ± 3% and +58% ± 2%, respectively.

Overall, the percentages of inclusion of the value "0" were quite close to their expected values (95%), which demonstrates the small bias of the method. The method was even unbiased when the QTLs were as close as 5 cM to a chromosome end, because 95.8% of the 95% confidence intervals contained the value "0" for a QTL with a heritability of 10% with the dense marker coverage and 94% with the sparse coverage (configuration 3 in Table 1). This last result contrasts with the properties of the confidence interval on a QTL's position, calculated using the same selective bootstrap, in which case it is anticonservative when the QTL is close to a chromosome end, as demonstrated by Lebreton and Visscher (1998). Only when the heritabilities of the two QTLs tested differed markedly, as in configuration 4 in Table 1, was the estimate of the "distance" significantly biased—by about 5 cM—for both marker densities. Even then, there was not a big discrepancy between the observed and the expected percentages of inclusions with the tight marker coverage. However, the bias was more important with the sparse coverage.

Selective *vs.* nonselective bootstrap: In Table 1, the selective bootstrap was also compared to the nonselective one. The 95% confidence intervals obtained with the nonselective bootstrap appeared consistently conservative, even very conservative when the QTLs were centrally situated (99.2%) in configuration 1. This implies that the hypothesis of pleiotropy is rejected too infrequently and therefore that the power of the test of close linkage *vs.* pleiotropy is not maximized with the nonselective scheme. The only configuration for which the confidence intervals were not too large was one in which two QTLs of high heritability (20%) were simulated near a telomere. The method was rather anticonservative, and the percentage of inclusion did not differ much from that obtained using the selective bootstrap. Average confidence interval widths on the estimated distance between the two tested QTLs vary from more than 210 cM in the least powerful design to ~90 cM in the most powerful of the simulated set, with the nonselective scheme. They can be reduced to about 125 and 67 cM, respectively, with the selective scheme. These confidence interval widths were estimated within confidence margins of ±1 cM only. Thus, the greatest reduction in confidence interval size is achieved for the least powerful designs, which makes sense, because a greater proportion of bootstrap samples were then rejected with the selective bootstrap.

Figure 2 illustrates the effect of the selection step in the bootstrap resampling, where two QTLs at different positions were simulated. It appears that the selection, in addition to reducing the bias of the test, quite dramatically increases its power, especially when the heritability is low. However, heritabilities as low as 10% offer limited resolution in the population we studied because, even when the QTLs are 180 cM apart, the power of the selective bootstrap test reached only 88.4%.

Effect of the marker density on the power of the test: In Figure 3, three marker densities were tested for two population sizes—100 and 200 genotypes. The results demonstrated that a higher marker density significantly increased the resolution of the test for both sizes. For a population size of 100 genotypes, a difference of 10–15% in power of the test was consistently observed between a sparse coverage of one marker every 20 cM and a saturated coverage of one marker every centimorgan, over the different distances between the QTLs. With the higher QTL detection power provided by the larger population size of 200, an even larger difference in QTL separation power was observed between the different marker densities.

Effect of the environmental correlation between the traits on the power of the test: In Figure 4, the results

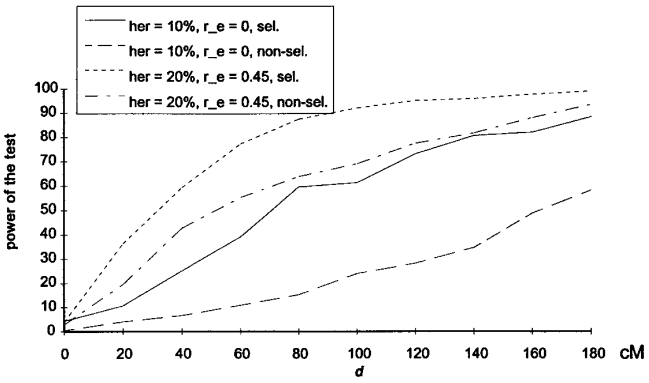

Figure 2.—Comparison of the power of the test with the selective and the nonselective bootstrap methods, across a range of distances (*d*) between the two QTLs, for two different heritabilities. The genome map was saturated with one marker every centimorgan. The simulated population size was 100. The chromosome length was 180 cM. Powers were estimated over 500 replicates for each configuration. r_e is the environmental correlation between the traits.

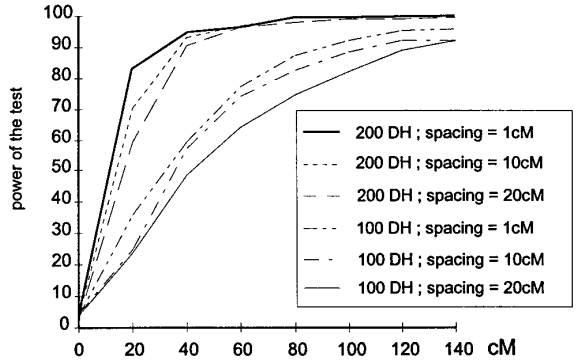

Figure 3.—Comparison of the power of the test across three marker densities and two population sizes, over a range of distances between the two QTLs. Only the selective bootstrap was applied. The powers were estimated over 500 replicates for each configuration. The simulated population sizes were 100 and 200, and the correlation coefficient between the residuals was 0.45. The chromosome length was 180 cM. Both trait heritabilities were 0.2.

of a series of simulations, where different degrees of environmental correlation were combined with different marker spacings, are plotted. The population size was 100, the trait heritability was 0.2, and the effect of the environmental correlation was symmetrical around the value zero. The power of the test increased with a high absolute value of the environmental correlation (r). The effect of an increase in r was more pronounced for tight marker coverage than for loose marker coverage. Thus, with one marker every 20 cM, we observed an increase of only 5% in the power of the test between $r = 0$ and $r = 0.9$, whereas with one marker every centimorgan the increase was 20%. This increase was explained by a smaller confidence interval on the estimate of d, the distance between the two QTLs, when r had a high absolute value. For $r = 0$, the average width of the 95% confidence interval of d was 77 cM but only 64 cM for $r = 0.9$ (results not shown).

Results presented in Figure 5 explore the effect of the environmental correlation in a less powerful design. The simulations followed the same protocol as that in Figure 4, and the population size was also the same, but the heritabilities of both traits were only 0.1. The maximum absolute value of r was also increased to 0.95 to accentuate any possible effect of this parameter. A very different pattern for the effect of r was observed. First, it was very asymmetrical around the value zero. Negative values of r dramatically increased the power of the test, unlike positive values. Surprisingly, the lower density of one marker every 20 cM provided a slightly better resolution except when $r = 0.45$. The average widths of the 95% confidence interval of d varied significantly between the different simulated configurations, from 112 ± 5 cM when $r = 0$ to 97 ± 6 cM when $r = -0.95$, for a marker spacing of 20 cM. Besides the differences in the size of the confidence interval of d, a bias in the estimate of d seemed to further exacerbate the difference in power associated with r. Thus, in Table 2, this bias was studied further. The average observed value of d was calculated over 2000 retained replicates and for the same range of genetic configurations as that simulated in Figures 4 and 5. Biases were significant only with the lower simulated trait heritability of 0.1 and larger with the lower marker density of one every 20 cM. A negative value for r tended to bias the estimate of d upward (+7.2 cM for $r = -0.95$ and one marker every 20 cM) and a positive value, downward (−6.9 cM for $r = 0.95$ and one marker every 20 cM). Likewise,

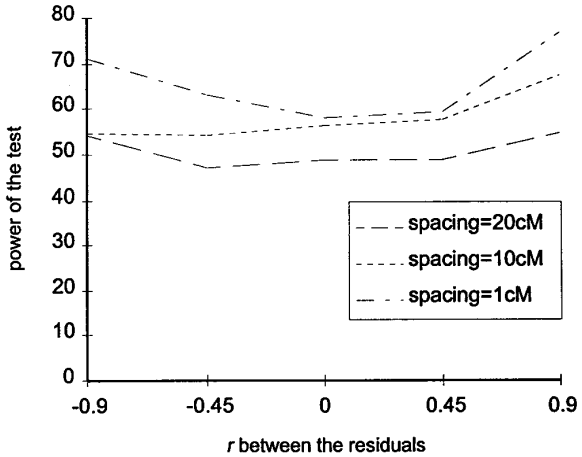

Figure 4.—Comparison of the power of the test across five values for the environmental correlation and three marker densities, for a distance of 40 cM between the two tested QTLs and trait heritabilities of 0.2. Only the selective bootstrap was applied. The powers were estimated over 500 replicates for each configuration. The chromosome length was 180 cM. The simulated population size was 100.

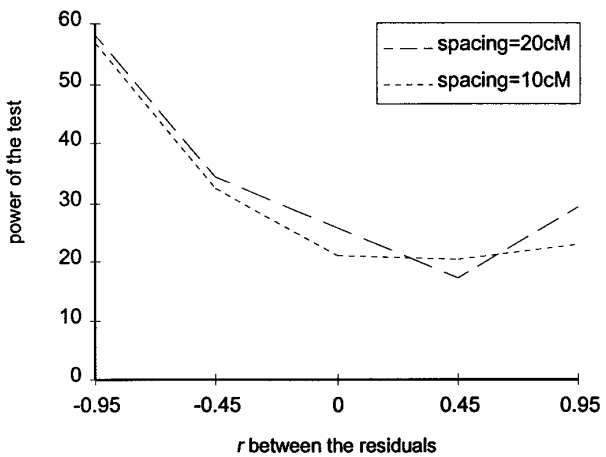

Figure 5.—Comparison of the power of the test across five values for the environmental correlation and three marker densities, for a distance of 40 cM between the two tested QTLs and trait heritabilities of 0.1. Only the selective bootstrap was applied. The powers were estimated over 500 replicates for each configuration. The chromosome length was 180 cM. The simulated population size was 100.

TABLE 2

Average estimated distance between QTL for trait B and QTL for trait A for linked QTLs

h^2	r_e	Marker spacing (cM)	D	σ_D	r_{pos}
0.2	−0.9	10	39.8	0.30	0.011 NS
		20	40.6	0.36	−0.014 NS
	0.0	10	39.7	0.30	0.0
		20	39.6	0.36	0.0
	+0.9	10	39.7	0.30	−0.027 NS
		20	38.6	0.34	−0.02 NS
0.1	−0.95	10	43.6	0.52	−0.146***
		20	47.2	0.55	−0.093***
	0.0	10	37.8	0.57	0.0
		20	39.2	0.59	0.0
	+0.95	10	36.0	0.49	0.242***
		20	33.1	0.53	0.205***

h^2, heritability of the trait; r_e, simulated correlation coefficient between the residuals; D, average estimated distance between the QTL for trait A and the QTL for trait B; σ_D, standard error of the mean; r_{pos}, correlation coefficient between QTL position estimates; NS, nonsignificant correlation; ***, significant at a risk of type I error <0.001. The simulated distance between the QTL for trait A and the QTL for trait B was 40 cM.

the correlation coefficients between the QTL position estimates were only significant with the lower simulated trait heritability of 0.1, but with both marker densities. A negative correlation between the residuals of the phenotypic trait values generated a negative correlation between the QTL position estimates and vice versa.

Behavior of the method in more complex genetic situations: The power and the bias of the test were also investigated in more complex configurations of several QTLs per trait. The model was three QTLs per trait, on a 160-cM-long chromosome. The QTLs were of equal effect and linked in repulsion. Both trait heritabilities were 0.6, and the residuals were correlated with $r = 0.2$. The population was made of 150 doubled haploid lines. One thousand replicates were simulated. The results are summarized in Table 3. QTLs 1, 3, and 5 affected trait 1, whereas QTLs 2, 4, and 5 affected trait 2. Thus, QTL 5 was pleiotropic by definition, and we had two testable pairs of QTLs: QTLs 1 and 2 in one region of the chromosome and QTLs 3 and 4 in another region. The QTLs were 13 cM apart for the first pair and only 8 cM apart for the second pair. This design was inspired partly by that of Jiang and Zeng (1995), although they simulated F_2 populations. The simulated chromosome was covered with only one marker every 20 cM to assess the robustness of the test in a realistic configuration. Both testable pairs of QTLs in close linkage were situated in the same marker interval. QTL 5 was the only QTL in its interval and was separated from the nearest one by one empty marker interval. Three positions within this interval were simulated for QTL 5. The estimate of the distance between QTL 1 and QTL 2 was very slightly biased upward by 1.6 cM and that between QTL 3 and QTL 4 was biased downward by ~4–5.6 cM, depending on the simulated position of QTL 5. The null hypothesis of pleiotropy, for QTL 5, was rejected with varying frequencies, from 3.6%, when the QTL was simulated in the middle of the interval, to 9.8% when it was simulated at only 2 cM away from the flanking marker on the left. Simulations with the saturated map showed biases in the same directions as those observed with one marker every 20 cM for the three positions of QTL 5, respectively (results not shown). This indicates not only that the biases observed for QTL 5 are due to the within-marker interval bias on QTL position estimates but also that the linkage with other strong effect QTLs differing for the two tested QTLs tends to affect the estimated distribution of d and the conservativeness of the test. The power of the test was relatively low for the other two pairs of nonpleiotropic QTLs. It ranged from only 4.2 to 7.2%, for QTL 3 and QTL 4, which were 8 cM apart, depending on the simulated position of the linked QTL 5. The power of the test for QTL 1 and QTL 2 was not affected by the position of QTL 5 and was ~44%.

Application of the test to real data: There was a significant difference between the parents for the length of leaf five. For Chinese Spring the value was 338.0 mm on average and for SQ1 was 214.5 mm. The distribution of this trait among the recombinant genotypes did not differ significantly from normality. The intensity of the stress generated by Na_2SO_4 reduced the number of spikes to fewer than one per plant. SQ1 produced on average one spike per plant, whereas Chinese Spring did not produce any spikes in any plant. Twenty-three genotypes did not produce any spikes at all. This generated as many missing data for the average spike length trait. However, the distribution of the mean values of the remaining genotypes did not differ significantly from normality.

Segregation distortion in the marker genotypes was observed on 11 chromosomes, representing every homoeologous group except group 6. In particular, chromosome *7A* showed some segregation distortion in several regions. Markers at the telomeric end of the short arm of chromosome *5A*, *i.e.*, the two leftmost markers on our map, showed significant segregation distortion ($P < 0.001$) in favor of the Chinese Spring-type alleles in a 72:28 ratio. A relatively good correlation was observed between the lines' phenotypic means for the two traits ($R^2 = 0.3$, results not shown).

The detected QTLs for the two traits are listed in Table 4. The 1000 permutations determined an *F*-significance threshold of 13.35 for spike length and 12.87 for leaf length. Four QTLs were detected for spike length and these accounted for 58% of the phenotypic variance among the line means. Two QTLs mapped on chromosome *5A*. The first QTL was quite strongly

TABLE 3

Parameters of QTL positions and effects and their estimates for data sets with three QTLs affecting each trait

	QTL1	QTL2	QTL3	QTL4	QTL5
			Parameters		
Position (cM)	23	36	81	89	127
					122
					130
Distance (cM)		13		8	0
Effect					
a1	1		−1		1
a2		−1		1	−1
			Estimates		
Distance (cM)		14.67 ± 0.66		4.14 ± 0.70	0.07 ± 0.5
				3.90 ± 0.60	−0.11 ± 0.3
				2.52 ± 0.66	2.44 ± 0.9
Effect (over replicates)					
a1	1.0 ± 0.02		−0.95 ± 0.02		1.36 ± 0.06
			−0.95 ± 0.02		0.98 ± 0.02
			−0.95 ± 0.02		1.54 ± 0.06
a2		−1.06 ± 0.03		1.48 ± 0.06	−1.63 ± 0.06
				1.52 ± 0.06	−1.5 ± 0.06
				1.32 ± 0.06	−1.45 ± 0.03
Effect (over bootstraps)					
a1	1.15 ± 0.02		−1.02 ± 0.02		1.36 ± 0.04
			−1.06 ± 0.02		1.10 ± 0.02
			−1.03 ± 0.02		1.47 ± 0.04
a2		−1.19 ± 0.02		1.47 ± 0.04	−1.58 ± 0.04
				1.54 ± 0.04	−1.54 ± 0.04
				1.37 ± 0.04	−1.45 ± 0.04
Power		0.444 ± 0.031		0.066 ± 0.015	0.054 ± 0.013
				0.076 ± 0.016	0.098 ± 0.018
				0.042 ± 0.012	0.022 ± 0.004

Position (cM), position of the QTLs in centimorgans from the leftmost marker, on the unique chromosome; Distance (cM), difference (QTL position trait 2 − QTL position trait 1) for a pair of QTLs testable for pleiotropy *vs.* close linkage; a1, additive effect of the QTLs affecting trait 1, with its 95% confidence limits; a2, additive effect of the QTLs affecting trait 2, with its 95% confidence limits. Power, for QTLs 1–4: power of the test of close linkage *vs.* pleiotropy, *i.e.*, frequency at which the null hypothesis of pleiotropy is rejected for a nominal risk of type I error of 5%; for QTL 5, percentage of type I error for a nominal risk of 5%.

The trait heritabilities are 0.6, the population size is 150, and the environmental correlation coefficient between the traits is 0.2. The chromosome is 160 cM long and covered with one marker every 20 cM. For QTL 5, three positions were simulated, and the three respective percentages of type I errors are listed bottom right. The corresponding powers of the test for QTLs 3 and 4 are also detailed. QTLs 1, 3, and 5 affect trait 1. QTLs 2, 4, and 5 affect trait 2. QTL 5 is therefore the only pleiotropic QTL.

significant ($F = 44.8$) and was estimated at 168.8 cM from the leftmost marker on our map. It is close to *Vrn1*, which mapped at 179.1 cM. The empirical distribution of its position estimate, using the selective bootstrap procedure described in Lebreton and Visscher (1998), did not allow us to reject *Vrn1* as its candidate gene. One would have had to adopt a 26% type I error risk for this rejection (results not shown elsewhere). The second QTL on *5A* for this trait mapped in a poorly covered region of the chromosome at position 229.0 cM in a 34.2-cM-wide marker interval and was linked in repulsion with the first QTL. *Vrn1* could be rejected as its candidate gene with a 1% type I error risk, again using Lebreton and Visscher's (1998) confidence intervals.

Only one QTL was detected for leaf length in the genome scan and this was at 202.6 cM on *5A*, ~20 cM from *Vrn1*, which required too high a risk of type I error (7%) to be rejected as its candidate gene. Figure 6 shows the comparative empirical distributions of the first QTL's position for the spike length and that of the QTL for the leaf length on *5A*. Applying our test of close linkage *vs.* pleiotropy to these two QTLs that mapped near *Vrn1* showed that the hypothesis of pleiotropy could be rejected with a stringency that allowed a 1% type I error risk. Figure 7 shows the empirical distribution of *D*, and zero clearly appears as an outlier to the distribution. Figure 8 shows that the correlation between the two QTLs' estimated positions was not at

TABLE 4

List of the QTLs detected for the average spike length and the leaf length characters

Chromosome	QTL no.	Position	Effect	*F*
		Average spike length (adjusted $R^2 = 0.58$)		
5a	1	168.8	0.59	44.80
5a	2	229.0	−0.60	32.56
5b	1	0.7	0.31	19.81
7a	1	414.39[a]	0.27	16.41
		Leaf length (adjusted $R^2 = 0.34$)		
5a	1	202.6	−39.08	48.15

Chromosome, name of the chromosome on which the QTL was detected; QTL no., order of the QTL on the particular chromosome; position, estimated distance in centimorgans from the leftmost marker on the chromosome; effect, estimated effect of the QTL; *F*, *F*-statistic of the QTL.

[a] The total genetic length is likely to be grossly overestimated due to some wide gaps between several linking blocks on this chromosome.

all significant. The test was also applied to the second QTL on 5*A* for the spike length, despite the poor marker coverage in its proximity, and to the QTL for the leaf length. Figure 9 shows the comparative distributions of these two QTLs' position estimates. Figure 10 shows the empirical distribution of *D* for this test.

DISCUSSION

We have introduced a novel way of testing close linkage *vs.* pleiotropy. Our method is based upon the empirical distribution of the distance between QTLs mapped for different traits in the same chromosome region. The method is very easy to implement and does not make any assumptions about the distribution of data or test statistics. Over the range of configurations tested with one QTL per chromosome, the method presented little bias, in that under the null hypothesis of pleiotropy the proportion of nonrejection was close to its expected value.

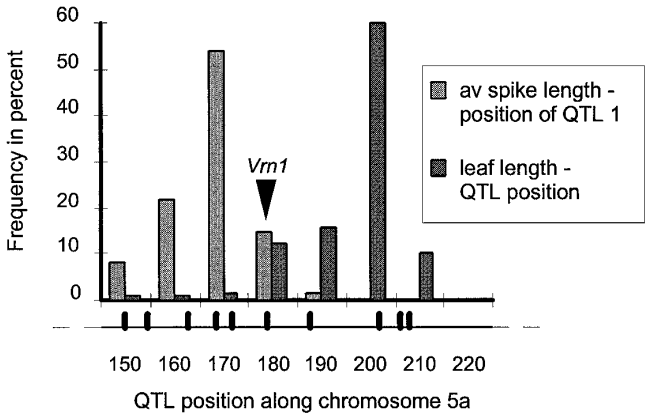

Figure 6.—Comparative empirical distributions of the estimated QTL positions obtained from 200 selective bootstraps applied on the original data, for the first pair of QTLs tested for close linkage *vs.* pleiotropy.

The QTL mapping method: Although the estimation of our QTL effect and position used the information from both flanking markers, the detection of the QTLs was carried out by selecting individual markers as opposed to testing pairs of adjacent markers as in the interval mapping or its multiple-linear approximations (Haley and Knott 1992). However, the consequent loss of power in our configurations with one marker every 20 cM is not very high, as shown by Knott and Haley (1992) and Rebai *et al.* (1995).

The selective bootstrap: The selection step inserted in the bootstrap resampling scheme was identical to that described in Lebreton and Visscher (1998). Because a significance threshold greater than zero was imposed to retain a QTL when analyzing an original data set, it was reasonable to apply the same threshold in order to follow the same procedure when analyzing the boot-

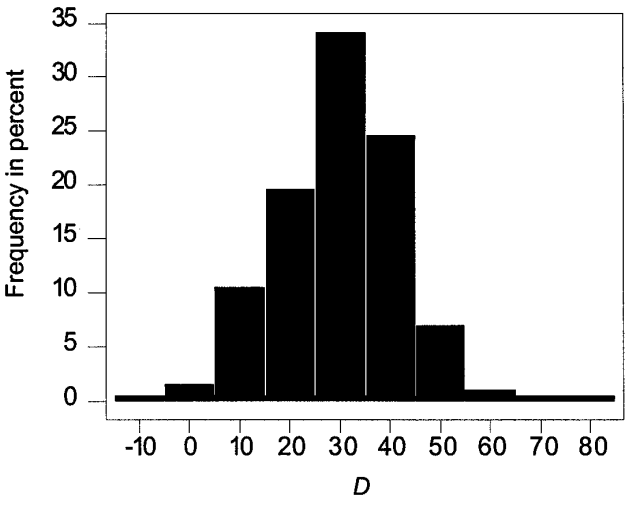

Figure 7.—Empirical distribution of the distance between the two QTLs of the first pair tested for close linkage *vs.* pleiotropy, obtained from 200 joint resamplings of the original data, using the selective bootstrap method. *D* is the estimated distance between the two QTLs.

strapped data sets to work out the empirical distribution of our *d* statistic (see Efron and Tibshirani 1993). This means that we only retained the bootstrap outcomes that showed the same QTLs for both traits. Selection was also imposed on the sign of the QTLs. This is because a confidence interval on a QTL's parameter is defined on the condition that this QTL exists. In other words, there is a genetic factor present somewhere along the chromosome, acting in a given direction on the trait. A QTL detected with an effect of opposite sign violates this condition and cannot logically be retained to contribute a value to the estimate of the statistic's empirical distribution. Nevertheless, the selection step generates an upward bias in the absolute value of the estimated QTL effect. As QTLs of higher effects have smaller confidence intervals on their position estimate, a prerequisite of the bootstrap is violated, namely that the statistic estimated from the bootstrapped data follows the same distribution law as that estimated from the original data set. Because we did not find the means to solve the issue analytically, we resorted to simulations.

In contrast to the situation in which the bootstrap resampling was applied on a single QTL's position, as in Lebreton and Visscher (1998) when the distance between two QTLs was resampled, in this study, no bias was observed when the pleiotropic QTL was situated near a telomere, provided that the percentages of variance that it accounts for are similar for both traits. In other words, on average, the distance is close to zero. Yet the individual estimated QTL's positions were biased toward the middle of the chromosome, as investigated by Hyne *et al.* (1995). The fact that two QTLs of similar significance, at the same position, are subject to similar biases may explain why their estimated distance remained close to the expected value of zero. Conversely, if a pleiotropic QTL explained a markedly different proportion of the variance of the two traits, the biases in the two estimated QTL positions would also be different and, as observed in our simulations, the estimated distance could differ significantly from zero. Even then, the conservativeness of the test did not seem overly decreased, which demonstrates the general robustness of the method.

Concerning the biases observed in the estimates of the QTL effects after the two selection stages, they are not solely an artificial consequence of the experimental protocol to assess our bootstrap method, but correspond to what would be observed when a set of real data is analyzed. Indeed, regarding the bias due to the first selection, if a pair of QTLs is to be tested for close linkage *vs.* pleiotropy, their observed effects would have to be significant for the QTLs to be detected in the first place. Imposing a significance threshold on the QTL effect for their detection involves a bias on their estimated effects, as investigated by Hyne *et al.* (1995). As for the bias due to the selection stage introduced in the bootstrap resampling, it grows with the percentage of outcomes rejected, which is inversely related to the size of the QTL effects.

Effect of the environmental correlation and of the marker spacing: We observed that the environmental correlation had a different pattern of effect, depending on the heritability of the traits. Thus, when the heritability was such that the experimental power was high (power of the test of close linkage *vs.* pleiotropy $>50\%$), then the resolution power of our test also increased significantly with a high absolute value of the environmental correlation *r*. This increase in resolution and the resolution of the test itself were greater with a higher marker density. When the traits' heritability was lower, a totally different pattern emerged. The effect of *r* was no longer symmetrical around the value zero, and it was the less dense marker density that provided a slightly higher resolution power of the test. When the heritabil-

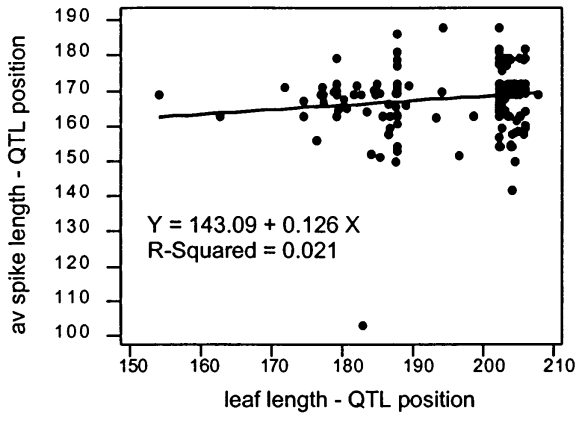

Figure 8.—Fitted regression between the estimated QTL positions of the first pair of QTLs tested for close linkage *vs.* pleiotropy, over the joint resamplings from the selective bootstrap method.

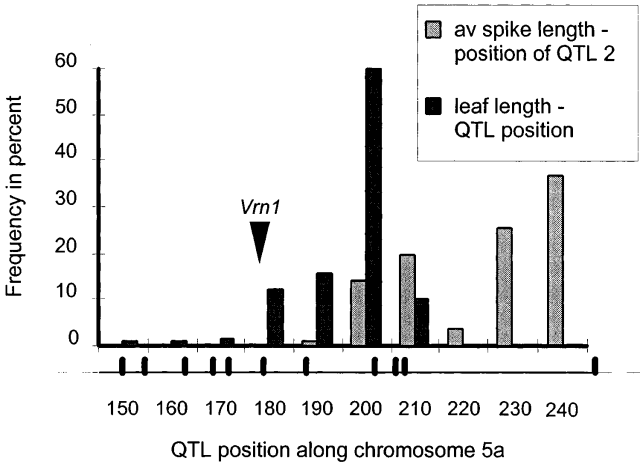

Figure 9.—Comparative empirical distributions of the estimated QTL positions obtained from 200 selective bootstraps applied on the original data, for the second pair of tested QTLs.

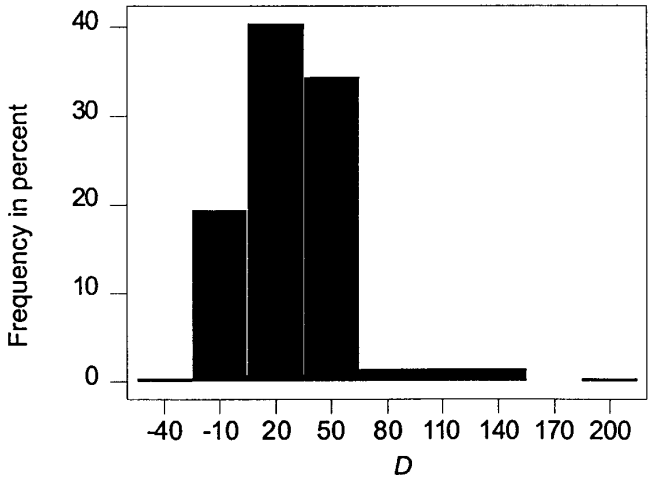

Figure 10.—Empirical distribution of the distance between the two tested QTLs of the second pair tested for close linkage *vs.* pleiotropy, obtained from 200 joint resamplings of the original data, using the selective bootstrap method. *D* is the estimated distance between the two QTLs.

ity was lower, the residuals represented a higher part of the phenotypic variation. Thus a strong correlation between the residuals such as those observed by Cheverud *et al.* (1997), for example, can have a dramatic effect on the resolution power as our simulations showed. The direction and magnitude of this effect are hard to predict intuitively. We observed that a positive *r* generated a positive correlation of the QTL position estimates between the two traits. This is expected to decrease the variance of *D* and therefore increase the power of the test. A negative *r* generates a negative correlation of the QTL position estimates between the two traits. This is expected to increase the variance of *D* and therefore decrease the power of the test. However, we also showed empirically that a positive *r* also generates a downward bias of *D*, which, if no other effects were present, would increase the apparent separation power of the test and vice versa. The overall effect is thus difficult to predict, and the simulations showed that the negative correlation between the residuals had the greatest effect in increasing the resolution; therefore, the bias on the QTL position estimates had the prevailing effect.

It also appears that the correlation between the environmental residuals of the two traits studied could not increase the resolution beyond a certain limit despite very high values of *r*. This is due to the fact that even if the environmental residuals were totally correlated, there would still be some discrepancy between the two traits' phenotypic scores—*i.e.*, the two vectors corresponding to these scores would not be colinear—due to the recombination between the two QTLs. In reality, of course, it is very likely that the traits would be determined by some additional genetic factors differing between the two traits. This would further decrease the gain in resolution obtainable from the correlation of the environmental residuals.

In the absence of bias on the QTL position estimates, the test of close linkage *vs.* pleiotropy showed very little bias itself. It also turned out to be robust in the presence of such biases on the QTL position estimates, whether they were due to differences between the two traits' heritabilities or to linkage with other QTLs at different distances as in the example presented in Table 3, provided that the marker density was high. However, when the marker density decreased, the test became more sensitive to these biases, including the within-marker interval bias on QTL position estimate.

Analysis of the wheat data set: The population sizes were small because data for a maximum of 96 doubled haploid lines were present in the data set. However, the high heritabilities of the traits studied compensated for the small sizes of our sample. The relevant chromosome span was covered with one marker every 6 cM on average for the first pair of QTLs tested for close linkage *vs.* pleiotropy, which mapped at some 34 cM away from each other. This allowed us to separate them with a maximum 1% type I error risk. The risk of type I error at which the second QTL for the spike length on chromosome *5A* can be separated from the QTL of leaf length is to be considered with caution because of the poor marker coverage around the former. Because our QTL mapping method fitted all the detectable QTLs on all the chromosomes, in the full model, no residual correlation of genetic origin was left. There was no correlation of environmental origin between the residuals because the data came from two separate experiments for the two traits. As a consequence, there was also no correlation between the reestimated QTL positions and no extra power to be gained from a joint resampling of the data in this particular case. More generally, for traits measured on the same individuals (as opposed to the case presented above), when implementing a QTL mapping method that includes all the detectable genetic effects in its full model, such as the composite interval mapping of Zeng (1993, 1994) or the multiple-QTL model of Jansen (1993), we also expect little extra power to be drawn from a joint resampling of the data if the correlation of environmental origin between the traits is low. However, in this configuration, the nonselective bootstrap method still has the merit of supplying an assumption-free nonparametric test that is robust in a variety of situations.

In conclusion, the selective nonparametric bootstrap offers a robust alternative, with little bias, to the LOD-based method to test close linkage *vs.* pleiotropy over an acceptable range of genetic configurations. It is computer intensive, but using our regression-based analysis the test is completed in, at most, a few minutes on any relatively new computer because only ~150–200 bootstraps are necessary. Another element of its attractiveness is the simplicity of its principle, hence its ease

of implementation in software using any QTL mapping method, provided the mapping method lends itself to the easy and quick estimation of the minimum number of QTLs along the chromosome of interest. It can be applied to the multiple-trait joint composite interval mapping of Jiang and Zeng (1995) through resampling of the estimated distance between the two QTLs of interest in the unconstrained full model, thereby taking advantage of the increased QTL detection power of their QTL mapping method. It is unlikely that our test brings any increase in separation power to the likelihood ratio-based test, but it is unlikely that any analytical method would bring the resolution of the test below 10–30 cM in common QTL mapping designs with realistic population sizes (Kearsey and Pooni 1996). To go below that resolution it will be necessary to resort to fine QTL mapping designs, such as advanced intercross lines or near-congenic lines, or to greatly increased population sizes.

We express our gratitude to M. J. Kearsey, J. W. Snape, and J. K. M. Brown for their helpful comments on the manuscript. C.S.H. is grateful to Ministry of Agriculture, Fisheries and Food (MAFF) and Biotechnology and Biological Sciences Research Council (BBSRC) for financial support. Finally, we are also grateful to the John Innes Foundation and Monsanto who funded the Ph.D. project of the corresponding author.

LITERATURE CITED

Basten, C. J., Z.-B. Zeng and B. S. Weir, 1996 QTLCartographer: a suite of programs for mapping quantitative trait loci, p. 108 in *Abstracts to Plant Genome IV*. Academic Press, San Diego.

Cadalen, T., C. Boeuf, S. Bernard and M. Bernard, 1997 An intervarietal molecular map in *Triticum aestivum* L. Em. Thell. and comparison with a map from a wide cross. Theor. Appl. Genet. **94:** 367–377.

Cheverud, J. M., E. J. Routman and D. J. Irschick, 1997 Pleiotropic effects of individual gene loci on mandibular morphology. Evolution **51:** 2006–2016.

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

Dempster, A. P., N. M. Laird and D. B. Rubin, 1977 Maximum likelihood from incomplete data via the EM algorithm (with discussion). J. R. Stat. Soc. Ser. **39:** 1–38.

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

Ellis, T. M. R., I. R. Philips and T. Lahaye, 1994 *Fortran 90 Programming*. Addison-Wesley, Wokingham, England.

Gale, M. D., M. D. Atkinson, C. N. Chinoy, R. L. Harcourt, J. Jia *et al.*, 1995 Genetic maps of hexaploid wheat, in *Proceedings of the Eighth International Wheat Genetics Symposium*, edited by Z. S. Li and Z. Y. Xin. China Agricultural Scientech Press, Beijing.

Galiba, G., S. A. Quarrie, J. Sutka, A. Morgounov and J. W. Snape, 1995 RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. Theor. Appl. Genet. **90:** 1174–1179.

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. Breeding **1:** 273–282.

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

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

Jiang, C., and Z.-B. Zeng, 1995 Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics **140:** 1111–1127.

Kearsey, M. J., and H. S. Pooni, 1996 *The Genetical Analysis of Quantitative Traits*. Chapman & Hall, London.

Knott, S. A., and C. S. Haley, 1992 Aspects of maximum likelihood methods for the mapping of quantitative trait loci in line crosses. Genet. Res. **60:** 139–151.

Korol, A. B., I. Y. Ronin and V. M. Kirzhner, 1995 Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics **140:** 1137–1147.

Lander, E. S., and D. Botstein, 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 184–199.

Lebreton, C. M., and P. M. Visscher, 1998 Empirical non-parametric bootstrap strategies in QTL mapping: conditioning on the genetic model. Genetics **148:** 525–535.

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.

Quarrie, S. A., M. Gulli, C. Calestani, A. Steed and N. Marmiroli, 1994 Location of a gene regulating drought-induced abscisic acid production on the long arm of chromosome 5A of wheat. Theor. Appl. Genet. **90:** 1174–1179.

Rebai, A., B. Goffinet and B. Mangin, 1995 Comparing power of different methods for QTL detection. Biometrics **51:** 87–99.

Rodolphe, F., and M. Lefort, 1993 A multi-marker model for detecting chromosomal segments displaying QTL activity. Genetics **134:** 1277–1288.

Ronin, Y. I., V. M. Kirzhner and A. B. Korol, 1995 Linkage between loci of quantitative traits and marker loci: multi-trait analysis with a single marker. Theor. Appl. Genet. **90:** 776–786.

Royston, J. P., 1982 Algorithm AS181: The W test for normality. Appl. Stat. **31:** 176–180.

Stam, P., 1991 Some aspects of QTL mapping. Proceedings of the Eighth Meeting of the Eucarpia Section Biometrics in Plant Breeding, Brno, Czechoslovakia, pp. 23–32.

Tinker, N. A., and D. E. Mather, 1995 MQTL: software for simplified composite interval mapping of QTL in multiple environments. J. Quant. Trait Loci. Avail. World Wide Web: http://probe.nalusda.gov:8000/otherdocs/jqtl/jqtl1995-02/jqtl16r2.html/.

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.

Whittaker, J. C., R. Thompson and P. M. Visscher, 1996 On the mapping of QTL by regression of phenotype on marker-type. Heredity **77:** 23–32.

Wright, A. J., and R. P. Mowers, 1994 Multiple regression for molecular-marker, quantitative trait data from large F2 populations. Theor. Appl. Genet. **89:** 305–312.

Xu, S., 1995 A comment on the simple regression method for interval mapping. Genetics **141:** 1657–1659.

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 of quantitative trait loci. Genetics **136:** 1457–1468.

Communicating editor: T. F. C. Mackay