Power and accuracy of QTL detection: simulation studies of one-QTL models

Gilles Charmet

Agronomie 20 (2000) 309–323

Original article

Power and accuracy of QTL detection: simulation studies of one-QTL models

Gilles CHARMET*

INRA, Station d'Amélioration des Plantes, 234 avenue du Brézet, 63039 Clermont-Ferrand, France

(Received 1 March 1999; accepted 6 January 2000)

Abstract – Nonparametric selective resampling procedures were used to investigate the effect of several factors on the power of quantative trait loci (QTL) detection, the bias and confidence intervals of their position, effect and heritability. The factors studied were population size, QTL position, heritability of the QTL and marker coverage, i.e., marker density and their regular versus random spacing. Confidence intervals obtained using either Normal approximation, the bias corrected and accelerated (BC_a) method and empirical bootstrap with 1 000 selected resamples were compared. The BC_a intervals were found to be very close to classic confidence intervals (CI) assuming normal distribution for sample sizes above 200, and to be slightly closer to empirical CI for small population sizes. The precision of the QTL position was found to be mostly affected by population size and heritability, and less by marker spacing, except in the case of sparse maps with irregular marker spacing. Bias in QTL position estimates can be high for small population sizes when QTL are located near the end of a chromosome, and, unexpectedly, selective bootstrap does not decrease this bias very much.

marker regression / bootstrap / detection power / QTL heritability/ linkage map

Résumé – La puissance et la précision de détection des QTL : études basées sur des méthodes de rééchantillonnage sélectif. On a utilisé des méthodes de rééchantillonnage sélectif pour étudier l'effet de plusieurs facteurs sur la puissance de détection des QTL (locus impliqué dans le déterminisme d'un caractère quantitatif), les biais et intervalles de confiances de leur position, effet et héritabilité (proportion de la variance du caractère expliquée par l'effet additif d'un QTL). Les facteurs étudiés étaient la taille de la population (haploïdes doublés), la position et l'héritabilité du QTL, la densité des marqueurs et leur espacement régulier ou au hasard. On a comparé les intervalles de confiance obtenus soit par l'approximation Normale, soit par la méthode BC_a (correction de biais et accélération), soit par bootstrap empirique avec 1 000 échantillons sélectionnés. Les intervalles obtenus par la méthode BC_a sont très proches de ceux obtenus avec l'approximation Normale pour les tailles de population supérieures à 200, et sont légèrement plus proches des intervalles empiriques pour les petites populations. La précision de localisation du QTL est surtout affectée par son héritabilité et la taille de la population, et relativement peu par l'espacement des marqueurs, sauf dans les cas extrêmes de cartes lâches

Communicated by Philippe Brabant (Gif-sur-Yvette, France)

* Correspondence and reprints
charmet@clermont.inra.fr

avec des marqueurs irrégulièrement espacés. Le biais dans l'estimation de la position d'un QTL peut être important quand celui-ci est proche de l'extrémité d'un chromosome et, de façon assez inattendue, le rééchantillonnage sélectif ne permet pas de réduire ce biais.

régression sur les marqueurs / rééchantillonnage / puissance de détection / héritabilité d'un QTL / carte de liaison génétique

1. Introduction

Since the advent of molecular marker linkage maps, quantitative trait loci or QTL have been searched for intensively in many crop species during the last decade. In addition to better knowledge of the genetic control of complex traits such as yield or quality, an improved response to selection is expected from the use of QTL-associated markers. Such QTL-marker associations can be used in two ways. The first is the use of a selection index including a molecular score, as first proposed [19] and further investigated [14, 26]. In this case, the power of QTL detection greatly influences the efficiency of such marker-assisted selection, as does the proportion of phenotypic variance explained by the QTL, further referred to as QTL heritability. The second way of using QTL-related markers in selection is their direct use as 'Mendelian' factors to be either transferred (e.g., by backcrossing) into a recipient elite genotype [35] or cumulated in a single genotype by an appropriate crossing programme [1, 4]. For such a use, it is important to avoid false positives, as a large number of segregating QTL will impede the efficiency of their transfer using a manageable population size [4], and to have a good coverage of the confidence intervals (CI) of QTL by markers to control their transfer [13].

It stems from these observations that estimates of detection power, rates of false positive (both associated with the individual and overall type-I error rate) and CI for QTL locations, heritabilities and effects are needed to optimise the use of QTL in plant breeding. However, most QTL mapping methods, based either on maximum likelihood [20] or on regression [12, 16, 24] do not lend themselves to a straightforward calculation of a CI for QTL location, effect or heritability. A widely used method of CI construction is the so-called LOD drop-off method [20], where the CI for QTL location is calculated by finding the location on either side of the estimated QTL location that corresponds to a decrease in the LOD score by 1 or 2 units. However, such CI are only valid asymptotically, and have been shown to be biased downwards for small- and medium-sized populations [23, 27]. For example, for a backcross of 200 individuals, the empirical probability that the 90% CI based on the 'one LOD drop-off' method contains the actual QTL location can be as low as 0.74 [23]. Hence this LOD drop-off method is no longer recommended in practice. Although [23] complex analytical formulae for CI have been derived for the case of two flanking markers, most alternative methods to calculate CI of QTL positions proposed so far have relied on simulation [7, 23]. Among them, the bootstrap method [8, 10] has been used for estimating CI of QTL locations and effects [15, 21, 34]. These authors investigated a wide range of experimental factors which may affect the width of CI, including population size, QTL heritability, QTL position, marker spacing, etc. Hyne et al. [15] assumed a Gaussian distribution of empirical distributions, which may not be correct for a limited number of bootstrap replicates (200). Alternatively, Visscher et al. [34] used an empirical CI, i.e., took the bottom and top 2.5th percentile of the ordered bootstrap distribution. Again, such a method may not perform the best, except when the number of bootstrap replicates is very high.

In this paper, we used the BC_a (bias corrected and accelerated) bootstrap method as proposed by DiCiccio and Efron [8] in comparison with Normal approximation and empirical bootstrap to improve the reliability of CI, and investigated further factors which would give a better fit for real situations, particularly that of irregular marker spacing. Our aim was also to summarise previous results on

detection power, estimation biases and CI length for QTL position, effect and heritability, and thereby give practical information to breeders when planning a QTL experiment.

2. Materials and methods

A linkage group of 160 cM length was generated with markers regularly spaced every 5 cM, which is usually considered as a 'dense' map, using the reverse of Haldane mapping function. The effect of QTL location was investigated by studying two positions: one near a chromosome end (2.5 cM), and one far to the extremity of the linkage group (52.5 cM), each position being in the middle of a marker interval, which is the least favourable position for QTL [35]. The additive QTL effect was arbitrarily set at a = 1, and a random, normally distributed noise with variance σ^2 was added to simulate phenotypic values of the trait, such that the ratio of additive variance explained by the QTL (a^2) on the total phenotypic variance $a^2 + \sigma^2$ was equal to a specified value called 'heritability of the QTL'.

A large population of 10 000 individuals was generated for each set of parameters studied, then B resamples of size N were drawn with replacement from this population to generate a bootstrap distribution. B was not chosen a priori; rather, a 'while' loop was used until 1 000 samples had been selected as having detected one QTL. This is not exactly the bootstrap procedure as defined in [10], which consists of resampling with replacement of the whole population observed. However, the drawback of using a single small population is that the initially generated set may be biased. Using a large population avoids this bias. Moreover, the empirical distribution of a given parameter obtained in this way is expected to follow bootstrap theory. Consequently, what we have called a CI in the following does not correspond exactly to the classical definition of a CI (i.e., an interval which covers the true value in 95% of cases), but rather reflects the expected sampling variation in the outcome of a QTL analysis. However, we have kept the term confidence interval (CI) for the sake of simplicity.

In each sampled dataset of size N, a single QTL model was fitted using the marker regression method described in [16], which provided least-squares estimates of two parameters: QTL position and effect. In the standard marker regression method, each linkage group is scanned from its origin to its end for the presence of a putative QTL. This means that the range of the QTL position is bounded by 0 and chromosome length and F. Hospital (pers. commun.) suggested that this could be a source of bias in QTL position estimate when the true position is close to a chromosome end. Therefore, for the case of a telomeric QTL, we used a modified marker regression programme, which allowed the scanning of a putative QTL to start at –50 cM, i.e., ahead of chromosome origin. A third parameter, QTL heritability, was estimated as the ratio of the square of the QTL effect on the phenotypic variance of the sample. Once a bootstrap distribution is obtained, it is straightforward to estimate its mean and standard deviation, and thus a CI can be obtained assuming a Gaussian distribution of the parameter (normal CI). However, if the distribution does not fit the Gaussian law, the true (i.e., 'hypothesis testing') CI also deviates from the normal CI, and its estimation is much more effort-demanding [8]. Therefore we compared three different confidence intervals: 1) the normal CI, i.e., mean $\pm$ z_0 standard deviation, z_0 being the p = 0.975 quantile of the normal distribution; 2) the empirical CI, i.e., the 26th and 975th values of the ordered distribution; and 3) the BC_a CI proposed by Efron and Tibshirani [10], which accounts for departure in skewness and kurtosis from the Gaussian distribution.

By definition, the BC_a endpoint is $\theta_{Bca} = G^{-1}\,\Phi\,[b_0 + (b_0 + z_\alpha)/(1 - ac\,(b_0 + z_\alpha)]$, where G is the cumulative distribution function (cdf) of the B bootstrap replications, Φ is the standard normal cdf, and z_α is $\Phi^{-1}(\alpha)$, b_0 is the bias correction and ac is the acceleration parameter, which measures how quickly the standard error is changing on the normalised scale. The computation of ac and b_0 can be found in [8]. It clearly appears that if ac and b_0 are zero, then

$\theta_{Bca} = G^{-1}(\alpha)$, the 100αth percentile of the bootstrap replications (see empirical CI). According to [8], the BC_a CI is second-order accurate, i.e., its precision increases as 1/N, N being the sample size, while the empirical CI is only first-order accurate (i.e., errors in matching go to zero at rate $1/(N^{1/2})$.

As the power of QTL detection depends largely on type-I error α, empirical detection power was established as the proportion of significant regression F values out of 1 000 trials. A range of individual type-I error risk was explored with N = 200 to achieve a global risk of 5%, as suggested in [5]. Further determinations of empirical detection power were then carried out at $\alpha = 0.001$, which yielded an average rate of false positives of 1–2%. Confidence intervals of QTL position, QTL effect and QTL heritability were then computed from the bootstrap trials which led to significant F values. This procedure is known as selective bootstrap and has been recently referred to as "conditioning to the genetic model" in [21]. The space of the parameters studied was defined as follows.

– For a dense, regular map of 25 markers, one every 5 cM, N ranged from 50 to 1 000, and h^2 from 0 (false positive) to 0.3. High heritability values (0.5, 0.75), such as those expected in narrow base crosses (e.g., between near isogenic lines used in fine mapping experiments), were also studied.

– For a sparse, irregular map and a medium-sized population of N = 200, various proportions from 100 to 25% of randomly chosen markers were used, in conjunction with heritability ranging from 0 to 0.3.

– The case of a regular, sparse map (one marker every 20 cM) was also considered to allow comparison with the sparse, irregular map.

All results are presented graphically, except those corresponding to very high heritabilities.

3. Results

3.1. Global type-I risk of false positives

The individual threshold for QTL testing was determined empirically by simulating datasets with no QTL. It clearly appears that an individual risk of 0.001 or less is necessary to keep the overall risk below 5%. As the detection power is positively related to α, the value of 0.001 appears to be a good compromise between detection power and the risk of declaring a false QTL, although it may be too conservative.

3.2. Detection power at $\alpha = 0.001$

Figure 1a shows the empirical power observed in 1 000 resamplings at various values of population size and QTL heritability for a centrally located QTL. A 90% threshold line was drawn. With very small populations (N = 50), it was exceeded only for highly heritable QTL ($h^2 > 0.30$), while very large populations (N = 1 000) allowed the detection of 'small' QTL ($h^2 = 0.025$). With the range of population sizes currently used, the critical heritability for which 90% power is obtained ranges from about 0.04 for N = 400 to 0.13 for N = 100. Substantial increases in detection power can thus be gained by increasing population size. 100 DH lines seems to be a critical limit below which only very large QTL can be detected, except by chance only.

For N = 200, detection power appears to be little affected by irregular marker spacing, at least for QTL with moderate heritability ($h^2 > 0.10$), as shown in Figure 1b.

Similar figures are given for a telomeric QTL in Figures 1c, d. While the power of detection of a telomeric QTL is of the same magnitude as that of a central QTL for a regular map, it decreases more rapidly with irregular marker spacing, particularly in the worst case (one marker picked up at random out of four). The most likely explanation is that the detection power should be very low when the only one marker lying at the left-hand side of the QTL is not retained in the map.

3.3. Bias and confidence intervals for QTL location

Table I summarises the extent of bias in QTL location, i.e., the difference between the average

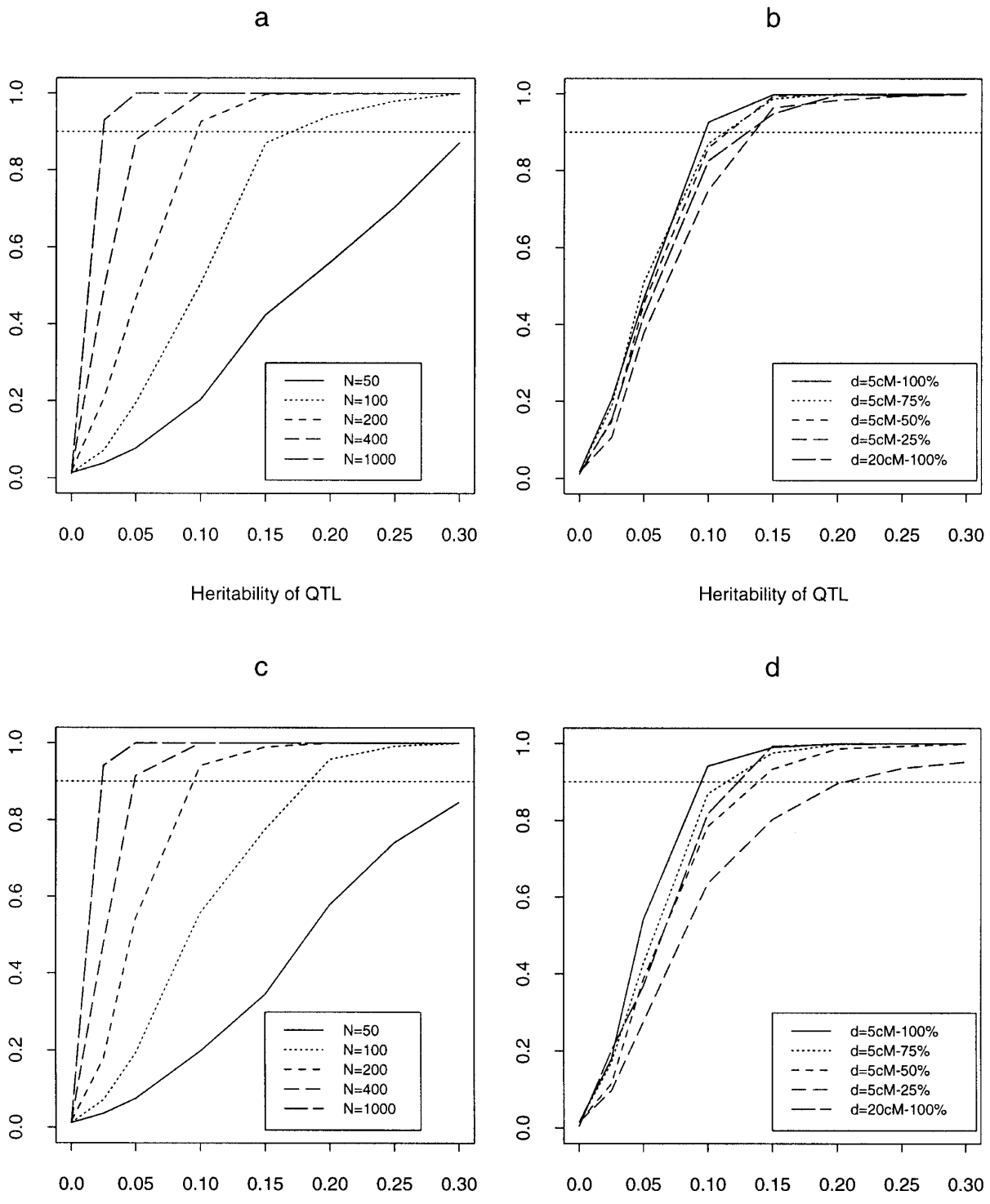

Figure 1. Power of QTL detection as a function of population size (N) and QTL heritability: **a:** for a centrally-located QTL and a dense, regular map; **b:** for a centrally-located QTL, N = 200 and uneven marker spacing (100, 75, 50 or 25% of the markers of the dense map randomly sampled); **c:** for a near telomeric-located QTL and a dense, regular map; **d:** for a near telomeric-located QTL, N = 200 and uneven marker spacing.

Table I. Bias in QTL position estimates, i.e., the difference between the average estimate over 1000 samples and the true position for a range of population sizes and QTL heritabilities. Values in italics in the second row are those obtained when negative estimates (up to –50 cM) are allowed.

QTL position			L = 2.5					L = 52.5		
N	50	100	200	400	1000	50	100	200	400	1000
$h^2 = 0.025$	47.82	38.68	26.75	12.13	6.17	12.10	7.77	–0.24	–1.64	–1.12
	28.44	*21.65*	*12.98*	*4.56*						
$h^2 = 0.05$	40.95	26.68	13.14	5.17	1.73	3.55	2.38	0.76	–0.19	0.74
	18.02	*15.46*	*8.80*							
$h^2 = 0.10$	25.08	12.12	5.18	2.10	0.89	1.57	2.81	1.49	0.98	0.79
	4.40	*2.67*	*0.59*							
$h^2 = 0.15$	13.28	6.17	2.51	0.21	–0.62	–0.58	0.02	–0.24	–0.70	–0.69
	3.65		*–0.82*							
$h^2 = 0.20$	14.48	5.84	1.87	1.93	–0.07	0.90	0.21	0.09	0.12	0.08
	–1.77	*–0.57*	*–1.08*							
$h^2 = 0.25$	8.59	2.62	1.07	0.69	0.27	–0.45	–0.68	–0.51	–0.36	–0.52
	–4.32	*–2.77*								
$h^2 = 0.30$	6.30	3.58	0.45	0.13	–0.28	0.14	0.46	–0.05	–0.10	0.02
	–6.19	*–5.44*								

position estimate over 1 000 samples with a dense, regular map, and the true (generated) QTL location. While this can be neglected for a centrally located QTL (52.5 cM), except for very low heritability, it may be as high as 40 cM for a telomeric QTL (2.5 cM). This bias increases as heritability and population size decrease (e.g., a bias of 10 cM towards the chromosome centre is found for N = 100 and $h^2 = 0.10$). When negative position is allowed for putative QTL, the bias in location estimate is strongly reduced and becomes more acceptable for $h^2 \geq 0.10$. The counterpart is a negative bias for small sample sizes and high heritabilities.

Tables II and III compare the 95% CI obtained with the three methods, namely empirical, normal approximation and BC_a-bootstrap, for a range of h^2 values and N = 200, for a centrally- and terminally-located QTL, respectively. Clearly, the normal CI is not appropriate for a telomeric QTL, as it leads to strongly negative lower bounds. The results of BC_a and empirical bootstrap are very similar. The BC_a method allows for asymmetric CI, which is not the case for normal CI, and it is less sensitive than empirical bootstrap to the effect of discrete scanning of linkage groups (every 1 cM). Although this asymmetry is not very marked, it may be more pronounced when the QTL is located towards the chromosome end [15], and therefore empirical or BC_a CI will be more accurate in any case.

The lengths of the 95% CI obtained from 1 000 BC_a bootstraps for a centrally located QTL are shown in Figure 2a. Threshold lines are drawn at 20 and 30 cM on the graph. It clearly appears that populations of limited size (N < 100) do not allow accurate estimation of QTL location, even for those that are highly heritable. Medium-sized populations (N = 200) make it possible to locate a QTL of intermediate heritability (0.10–0.15) within about 30 cM. This length is still rather large for QTL pyramiding through marker-assisted selection, as the probability of recombination between markers and QTL makes it necessary to use large populations to be able to manipulate several QTL at a time [4]. Reasonably narrow CI for QTL pyramiding can be obtained with N = 400 for intermediate heritability and for N = 1 000 (a very large population with current marker technology) for low heritability. However, even with large populations and high heritabilities such as those obtained

Table II. Comparison of different CI obtained for a 'centrally' located QTL (L = 52.5) from 1 000 resamples with N = 200 and various QTL heritabilities: lower and upper bounds (in brackets) and CI length (in italics), for normal approximation, bias correction and acceleration (BC_a), and empirical distribution.

N = 200	Bootstrap mean	Normal	BC_a	Empirical
$h^2 = 0.75$	52.54	47.46–57.62	47.30–57.46	47–58
		10.16	*10.16*	*11*
$h^2 = 0.50$	52.36	46.19–58.53	46.44–58.79	46–58
		12.34	*12.35*	*12*
$h^2 = 0.30$	52.44	43.48–61.40	43.76–61.68	44–62
		17.92	*17.92*	*18*
$h^2 = 0.25$	52.46	41.90–62.07	41.04–61.18	43–63
		20.17	*20.14*	*20*
$h^2 = 0.20$	52.58	41.97–63.20	41.72–62.94	42–63
		21.23	*21.22*	*21*
$h^2 = 0.15$	52.28	39.95–64.57	40.89–65.53	40–64
		24.62	*24.64*	*24*
$h^2 = 0.10$	53.94	35.65–72.33	36.06–72.74	36–73
		36.68	*36.68*	*47*
$h^2 = 0.05$	53.25	22.87–83.64	27.31–88.22	22–86
		60.77	*60.91*	*64*
$h^2 = 0.025$	54.90	3.27–106.54	12.16–115.96	0–120
		103.27	103.80	*120*

Table III. Comparison of different CI obtained for a QTL located near a chromosome end (L = 2.5) from 1 000 resamples with N = 200 and various QTL heritabilities: lower and upper bounds (in brackets) and CI length (in italics), for normal approximation, BC_a, and empirical distribution. Location values in italics on the second line are those obtained when negative estimates (up to –50 cM) are allowed.

N = 200	Bootstrap mean	Normal	BC_a	Empirical
$h^2 = 0.75$	3.21	–1.87–8.30	0.06–10.30	0–8
	(2.11)	*10.17*	*10.35*	*8*
$h^2 = 0.50$	3.18	–2.64–9.01	–0.30–11.44	0–11
	(1.09)	*11.65*	*11.74*	*11*
$h^2 = 0.30$	4.66	–3.63–14.85	–0.86–17.66	0–17
	(1.55)	*18.38*	*18.52*	*17*
$h^2 = 0.25$	4.26	–5.54–14.06	–1.32–18.47	0–19
	(1.57)	*19.60*	*19.79*	*19*
$h^2 = 0.20$	5.38	–5.78–16.53	–3.21–19.21	0–22
	(1.42)	*22.31*	*22.42*	*22*
$h^2 = 0.15$	5.23	–12.03–20.49	–2.50–31.36	0–32
	(0.68)	*32.52*	*33.86*	*32*
$h^2 = 0.10$	*7.68*	*–17.17–32.53*	–3.45–47.29	0–43
	(3.09)	*49.70*	*50.753*	*43*
$h^2 = 0.05$	*15.66*	*–40.34–71.66*	–2.91–112.27	0–115
	(11.30)	*112.0*	115.18	115
$h^2 = 0.025$	29.25	–43.74–102.24	–11.11–109.54	0–139
	(15.48)	*145.98*	*120.65*	*139*

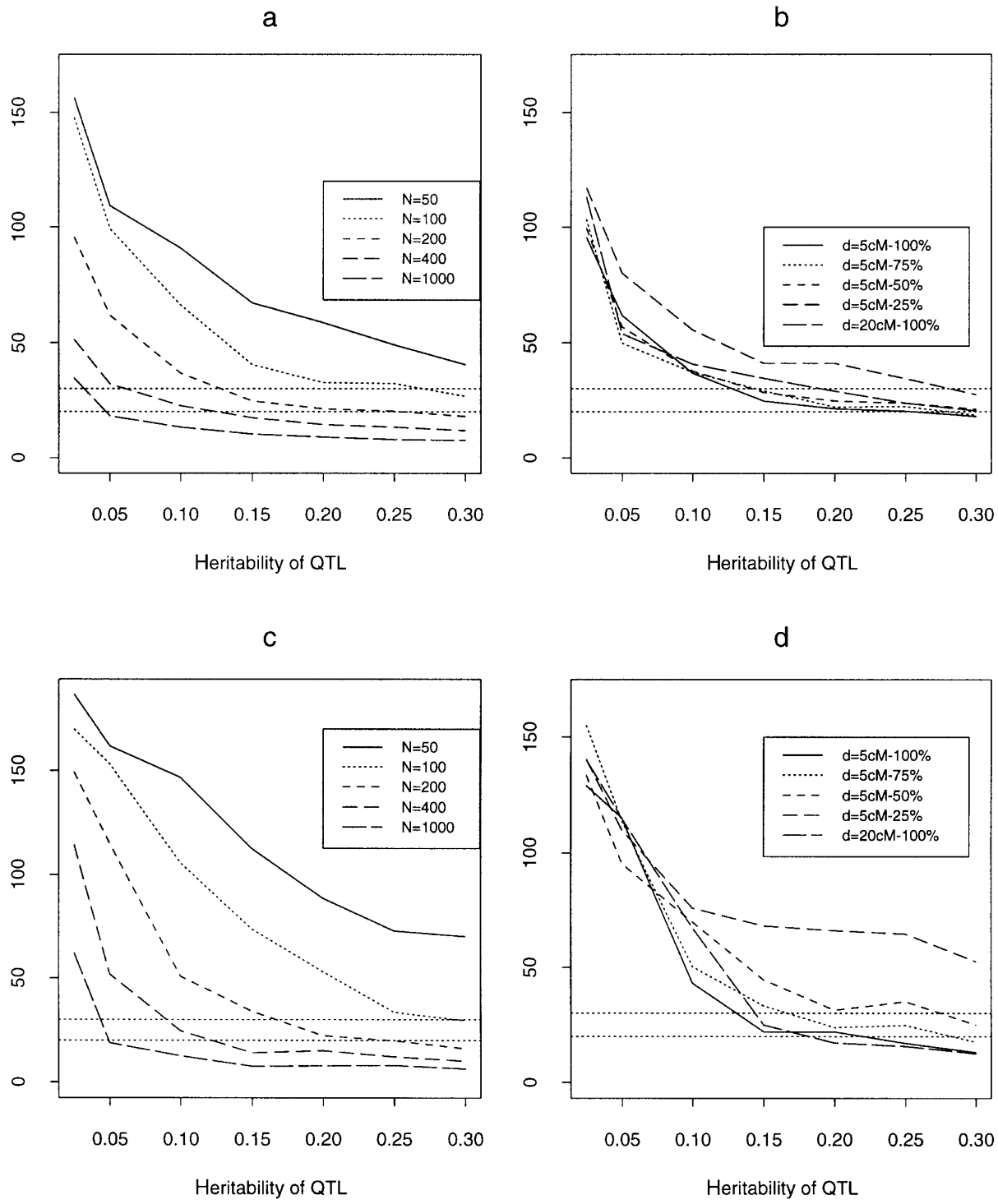

Figure 2. Length of CI by BC_a bootstrap for QTL position as a function of population size and QTL heritability: **a:** for a centrally-located QTL and a dense, regular map; **b:** for a centrally-located QTL, N = 200 and uneven marker spacing; **c:** for a near telomeric-located QTL and a dense, regular map; **d:** for a near telomeric-located QTL, N = 200 and uneven marker spacing.

in fine mapping designs, the accuracy of QTL location is far from the 1 cM precision required for chromosome walking and gene cloning (about 10 cM for $h^2 = 0.75$ and N = 1 000).

Figure 2b shows that the accuracy of the QTL position is little affected by irregular or wide marker spacing, except for the poorest case (an irregular map with an average interval length of 20 cM). However, at intermediate heritabilities, a substantial gain in detection power can be obtained by using a dense map (e.g., for $h^2 = 0.10$, p = 0.82 with a sparse map, 0.95 with a dense map).

Results for a telomeric QTL are given in Figures 2c, d. Regarding detection power, the CI of the position of a telomeric QTL is more affected than that of a central QTL by a very irregular spacing of markers.

3.4. CI of QTL effect and heritability

The main feature to be noticed in Figure 3 is that the accuracy of QTL effect estimates is very poor for QTL with low heritability, and decreases only slowly with increasing heritability: a CI width of 0.5 (for a simulated additive effect of 1) is reached only for a QTL heritability of > 0.15 with N = 400. The consequence is that prediction of genetic gain by cumulating favourable QTL alleles will be very inaccurate. Again, the effect of irregular marker spacing is very small, unless markers are very sparse (Fig. 3b, d).

The same remarks can be made for QTL heritability CI, shown in Figure 4a and c for a centrally-located and a telomeric QTL, respectively. Indeed, heritability has been estimated as the ratio of the square of the additive effect on the phenotypic variance. Its accuracy is thus expected to be poorer than that of the QTL effect, as both the numerator (its variance being four times that of the QTL effect) and the denominator are subjected to sampling errors. A lower CI bound of zero means that the experiment lacks power. It should be noticed that very high estimates of QTL heritability can be highly overestimated in small samples (N < 200), even for a simulated QTL with moderate heritability. Moreover, Figures 4b and d clearly show that heritability estimates are biased upwards for low heritabilities and small sample sizes. This must remind us that 'big' QTL obtained in real experiments with low- or medium-sized populations should be treated with caution, and validated in a separate experiment as often as possible. There are very small difference in CI estimates among regularly and irregularly spaced marker maps, except for very sparse ones (results not shown).

4. Discussion and conclusion

The power and accuracy of QTL detection have been investigated in several studies, either analytically or by means of simulations. The theoretical aspects of the use of analysis of variance (ANOVA) to detect linkages between marker locus and QTL have been reported [2, 11, 17, 18, 30–33]. These studies were based on the use of single markers, and thus the recombination rate between marker and QTL has to be taken into account, either explicitly [30], or implicitly [11]. More recently, it has been shown that initial theoretical computations based on ANOVA markedly underestimated detection power [3], because expectations of mean squares were wrong. The correction proposed fitted previous results based on simulation much better. The detection power of interval mapping methods, such as that proposed in [20], has been investigated by means of simulation [2, 27]. Rebai et al. [29] showed that interval mapping methods are only slightly more powerful than one-way ANOVA, at least for intervals < 30 cM.

Confidence intervals for QTL location and effect have been mostly explored through simulation studies. Mangin et al. [23] proposed an analytical method for constructing an unbiased CI of the location parameter, estimated by the maximum likelihood method. However, the difficulty of their method lies in the computation of the correct threshold for the maximum likelihood ratio test. Mangin and Goffinet [22] worked further on this method and proposed an approximation for the threshold. They also carried out a simulation to compare the length of their asymptotically similar

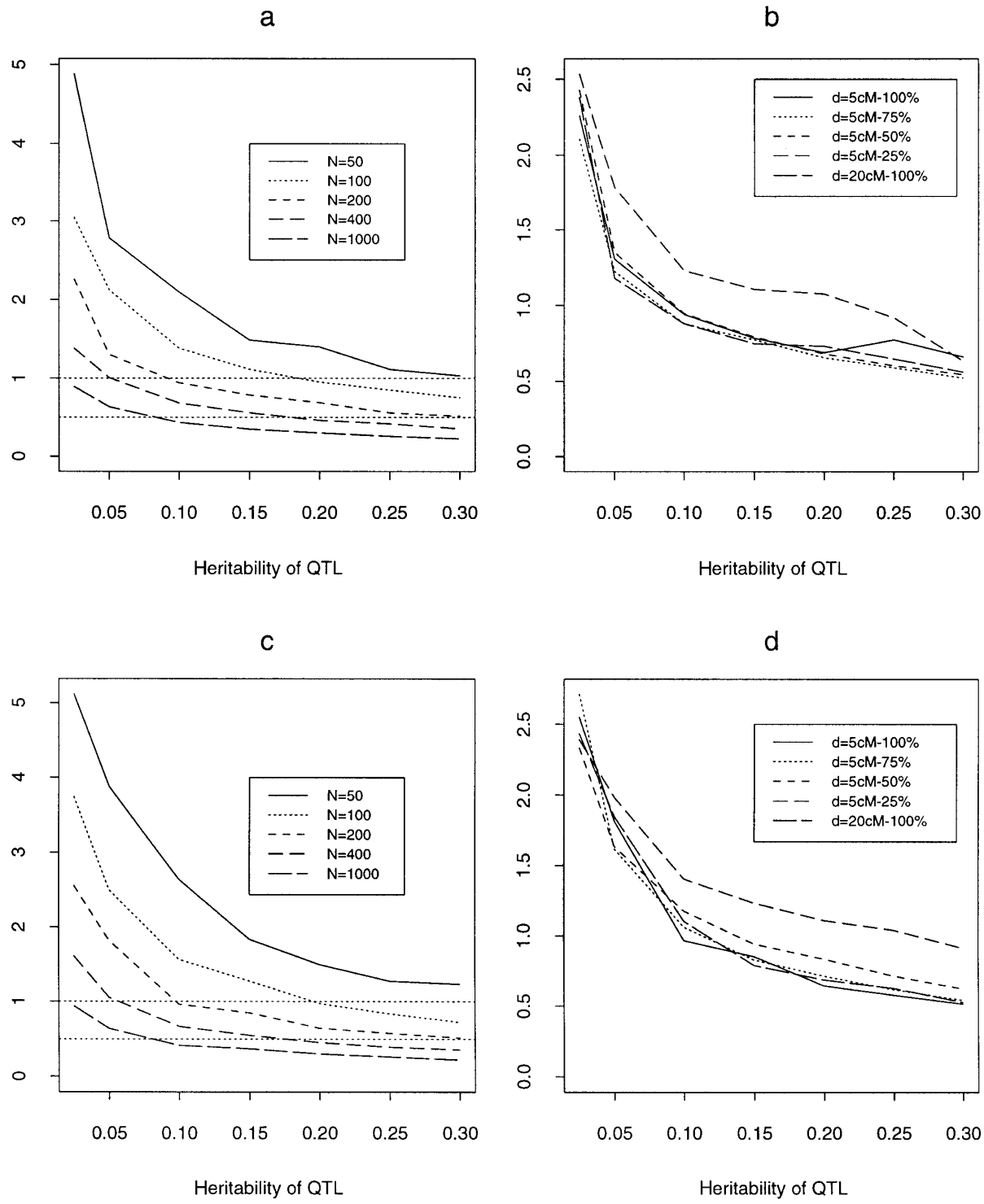

Figure 3. Length of CI by BC_a bootstrap for QTL effect as a function of population size and QTL heritability: **a:** for a centrally-located QTL and a dense, regular map; **b:** for a centrally-located QTL, N = 200 and uneven marker spacing; **c:** for a near telomeric-located QTL and a dense, regular map; **d:** for a near telomeric-located QTL, N = 200 and uneven marker spacing.

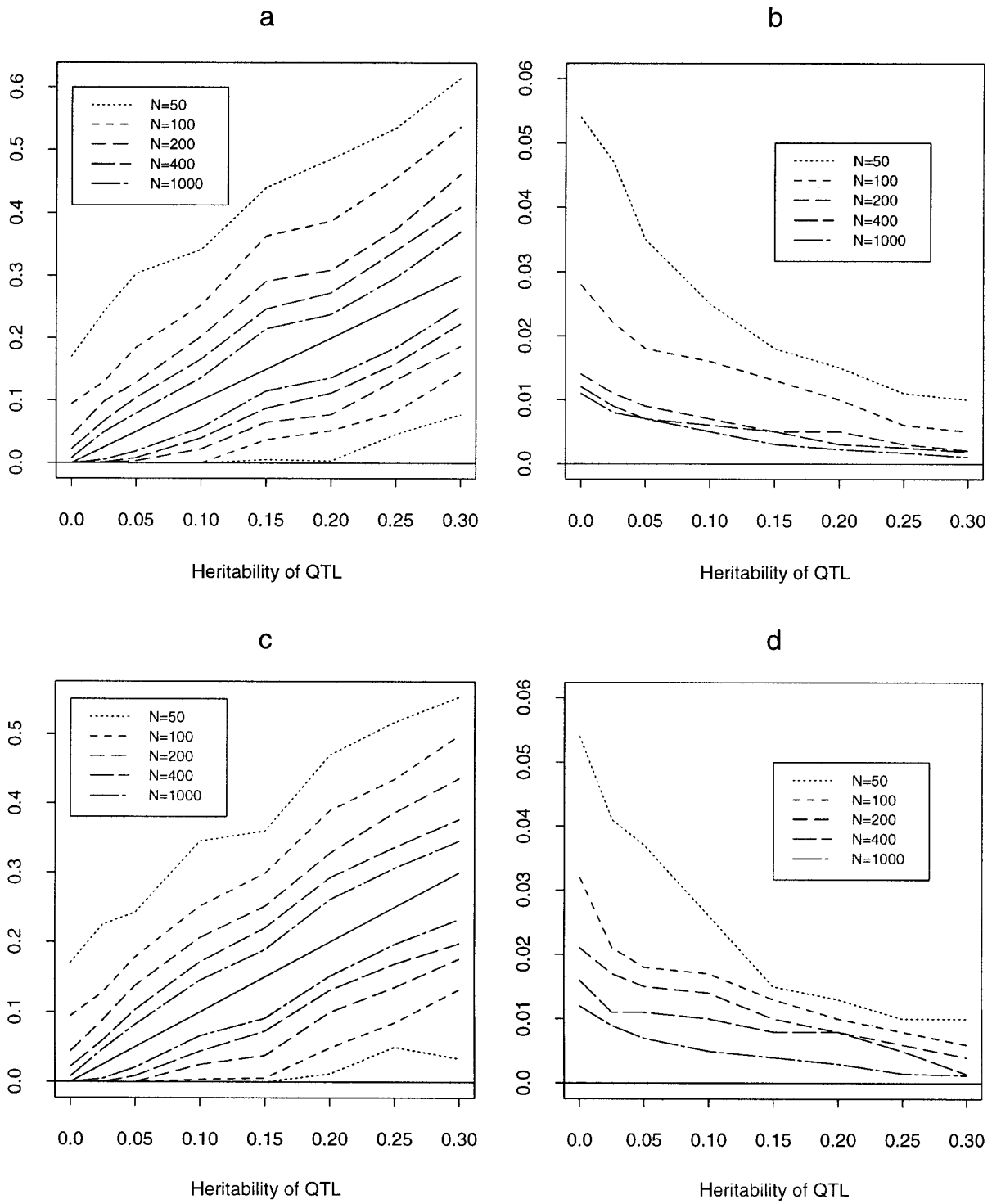

Figure 4. Lower and upper bound and average bias of QTL heritability estimate as a function of population size and QTL heritability: **a:** CI for a centrally-located QTL and a dense, regular map; **b:** bias for a centrally-located QTL and a dense, regular map; **c:** CI for a near telomeric-located QTL and a dense, regular map; **d:** bias for a near telomeric-located QTL and a dense, regular map.

CI to that of an empirical symmetrical one which had been previously proposed [7], and found that their CI could be half the length of the symmetrical one, particularly when detection power is low. Alternatively, bootstrap has been proposed [21, 34] as an intuitively simpler method, although requiring computer time. In a more recent study [36], several bootstrap procedures to construct CI in QTL mapping have been compared: the non-parametric method, as used in this study, was found to produce results close to expectation, while the parametric method performed poorly.

However, none of these studies considered the same type of progenies, and this makes the comparison of results difficult. Let us consider the proportion of phenotypic variance explained by a QTL, often called heritability of the QTL, which is appropriate when only an additive effect is assumed. For a given value of additive effect of allele substitution, a, and assuming no dominance effect, the heritability of the QTL is a^2/σ_p^2 ($\sigma_p^2 = \sigma_e^2 + a^2$) in a DH population, $a^2/2\sigma_p^2$ ($\sigma_p^2 = \sigma_e^2 + a^2/2$) in an F2 population and $a^2/4\sigma_p^2$ ($\sigma_p^2 = \sigma_e^2 + a^2/4$) in a BC population, where σ_p^2 is the phenotypic variance and σ_e^2 is the phenotypic variance not explained by the QTL. If the QTL under study was the only one affecting the trait, then σ_e^2 is the true non-genetic variance. Only in this case, and for a^2 small compared to σ_e^2, the heritability of the QTL is twice as high in an F2 population as in a BC population, and four times as high in a DH population, and one could conclude that an F2 population is roughly twice as powerful and a DH population nearly four times as powerful for detecting a certain QTL as a BC population. However, these two-fold (F2) or four-fold (DH) genetic variances hold for other QTL as well (if there is more than one QTL) [27]. Therefore, the phenotypic variance not explained by the QTL considered also increases substantially, even if the true non-genetic variance remains the same. Hence, the fraction of the variance explained by a given QTL in an F2 will be less than two-fold, and that of a DH population less than four-fold this fraction in a BC population. Thus, comparison of simulation results is very dependent on the parameters used in the simulations. For example, Carbonell et al. [2] simulated BC and DH populations of 250 individuals with six QTLs having various heritabilities. With these parameters for doubled haploids, they found a detection power of about 90% for QTL heritabilities as low as 0.05, while to obtain a similar power for backcrosses the heritability attributable to an individual QTL should be around 14%.

Moreover, this heritability attributable to an individual QTL may not be the appropriate parameter to be used in comparing results from the literature. Recently Dupuis and Siegmund [9] gave precise theoretical results to compare powers of QTL detection and lengths of CI for a dense map. Their formulae involve a non-centrality parameter, i.e., $\xi = \{N \ln (1 + a^2/\sigma_e^2)\}^{1/2}$ for doubled haploids and $\xi = \{N \ln (1 + a^2/4\sigma_e^2)\}^{1/2}$ for backcrosses. So it could be suggested that results for DH should be compared with those for BC using these parameter functions instead of the fraction of the total variance explained by the QTL. However, it is often not possible to estimate this non-centrality parameter from published data. Therefore, we will restrict the comparison of our results to those obtained with simulated DH populations. We have deliberately chosen this type of progeny, which has become a favourite material in genetic studies, since it can be produced quickly in a range of crops and provide 'fixed' genotypes, which are suited for replicated trials.

The study in [15] compared the accuracy of QTL detection in both F2 and DH populations, and is the only work which used the marker regression approach [16], as in the present study. As expected, the results of Hyne et al., based on normal distribution assumptions and only 100 resamplings, are very similar to ours for a centrally-located QTL. However, we have simulated a wider range of conditions, particularly for population size, and cannot fully agree with their conclusions. Although there is little increase in accuracy from using 150 instead of 100 DH lines, as stated in [15], we report that significant improvement in accuracy could be gained by using, for example, N = 400 instead of

N = 100 to allow practical manipulation of QTL. Hyne et al. [15] also compared a dense (every 5 cM) versus sparse (every 20 cM) marker map, and found that both results were in good agreement. They also compared CI obtained by assuming fixed marker position (i.e., known without error) to those obtained by re-estimating marker position in each simulated dataset, and found few differences. We also carried out this comparison and can confirm their conclusion. Moreover, the observed trend is that slightly larger CI are obtained when using re-estimated marker positions. Last but not least, Hyne et al. [15] also studied the case in which a QTL is located near a chromosome end. They found a considerable bias in mean location estimates towards the middle of the chromosome, ranging from 4 cM for $h^2 = 0.10$ to up to 20 cM for $h^2 = 0.02$, for a population size of 300 F2 (or 150 DH) in both cases. This bias towards the centre is caused, according to Walling et al. [36] and Xu [37], by the detection of false positives uniformly distributed along the chromosome. Hyne et al. [15] considered all simulated subsets, whether or not the QTL was significant. One therefore expects that this bias should disappear by using selective bootstrap, i.e., using only those subsets where a significant QTL has been detected for constructing CI [21]. Results with a simulated QTL at 2.5 cM from the end shows that the bias still exists, although we performed selective bootstrap, and that it increases when the QTL effect is small, thus confirming the results of Hyne et al. [15]. Again this bias is not reduced, but rather further increased when using re-estimated marker position based on the samples instead of their 'true' locations. A possible explanation may be that selective bootstrap is not stringent enough, and that many false-positive QTL, randomly distributed, do escape the selection step when using small populations. More recently, Walling et al. [36] reported that the bias of position in QTL mapping varied within a marker interval, being positive in the middle and negative at the marker position, and noted the relative inefficiency of selective bootstrap in correcting this bias. Alternatively, allowing negative estimates of QTL locations led to a strong reduction of bias. Obviously, on simulated data, we know that negative estimates are unrealistic, and therefore we did not consider this possibility in a first step. But on real data it is rarely known whether the terminal markers are truly on telomeres, and the possibility of putative QTL outside the covered linkage group must be considered.

Most of the previous studies considered a regular, dense genetic map with one marker every 5 or 10 cM. Rebai et al. [29] tested a range of marker intervals, and found a loss in power of about 15% between d = 0 and d = 40 cM. We found a similar result (Fig. 1b). However, with specific heritability corresponding to the inflexion point of the power curve, the reduction in power may be greater, as observed from 1 000 simulations for $h^2 = 0.05$ and N = 200 ($p = 0.616$ for d = 5 cM and $p = 0.44$ for d = 20 cM). The effect of irregular marker spacing, which had never been reported before, is very low (e.g., for $h^2 = 0.05$ and N = 200 and for an average spacing of 20 cM, $p = 0.424$ for random distribution of markers, versus $p = 0.444$ for regular spacing).

The results of our simulations largely confirmed previously published results. Some of the discrepancies might be due to the use of different methods for QTL detection and mapping. Indeed, most methods are based on maximum likelihood and LOD scores, while we used linear regression because it is straightforward to programme and fast-running. However, Kearsey and Hyne [16] have shown that the estimates of QTL location and effect by marker regression are consistent and as reliable as those given by conventional interval mapping methods. It is therefore unlikely that our results would have been changed significantly by using another QTL analysis method. Population size and QTL heritability are the main factors affecting detection power, which could therefore be improved by reducing the residual variance. This can be achieved by controlling environmental variation by means of replications, or by controlling genetic variation caused by other QTL. Gallais and Rives [11] have already stressed this increase in detection power of small QTL when including previously detected larger QTL in the model. Another main feature is the relative inaccuracy of estimates of QTL location and effect, which can be

expected from current-sized experiments. This lack of power and accuracy has also been reported in real QTL analysis in maize [25]. Probably, for a preliminary scan of a segregant population, it would be more appropriate to use the term 'QT regions', rather than 'QT loci'. As already mentioned by Hyne et al. [15], there is little to be gained in accuracy by using a dense versus a sparse genetic map. Our results, obtained in a more realistic way (i.e., non-constant marker spacing), are very similar and show that an optimum average marker spacing could be estimated, depending on the relative costs of genotyping and trait evaluation, as suggested by Darvasi and Soller [6]. As a concluding remark, it should be stressed that even for highly heritable QTL, the CI for location should never fall below 1 cM, a threshold commonly proposed to achieve positional gene cloning. Other strategies such as fine mapping based on near isogenic lines [28] or the use of gene homology with fully sequenced model genomes should be preferred. However, QTL identified in medium-sized populations (N = 200–400) could be manipulated as statistical objects [1, 4]. The CI of interactive QTL will be dealt with in a subsequent paper.

Acknowledgements: The author wishes to thank Dr F. Vear for her assistance with the English language, and Dr F.X. Oury and N. Robert for helpful discussion and comments. I am also grateful to the two anonymous referees who gave very pertinent advice and largely contributed to the improvement of the manuscript, and also to Dr F. Hospital for his perspicacious suggestion of the possibility of QTL location outside the known linkage group for reducing bias.

References

[1] Berloo van R., Stam P., Marker-assisted selection in autogamous RIL populations: a simulation study, Theor. Appl. Genet. 96 (1998) 147–154.

[2] Carbonell A.E., Asins M.J., Balsega M., Balansard E., Cerig T.M., Power studies in the estimation of genetic parameters and the localization of quantitative trait loci for backcross and doubled haploid populations, Theor. Appl. Genet. 86 (1993) 411–416.

[3] Charcosset A., Gallais A., Estimation of the contribution of quantitative trait loci (QTL) to the variance of a quantitative trait by means of genetic markers, Theor. Appl. Genet. 93 (1996) 1193–1201.

[4] Charmet G., Robert N., Perretant M.R., Gay G., Sourdille P., Groos C., Bernard S., Bernard M., Marker assisted recurrent selection for cumulating additive and interactive QTL in recombinant inbred lines, Theor. Appl. Genet. (in press).

[5] Churchill G.A., Doerge R.W., Empirical threshold values for quantitative trait mapping, Genetics 138 (1994) 963–971.

[6] Darvasi A., Soller M., Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci, Theor. Appl. Genet. 89 (1994) 351–357.

[7] Darvasi A., Weinreb A., Minke K., Weller J.I., Soller M., Detecting marker-QTL linkage and estimating QTL gene effect and map location using a sturated genetic map, Genetics 134 (1993) 943–951.

[8] DiCiccio T.J., Efron B., Bootstrap confidence intervals, Stat. Sci. 11 (1996) 189–228.

[9] Dupuis J., Siegmund D., Statistical methods for mapping quantitative trait loci from a dense set of markers, Genetics 151 (1999) 373–386.

[10] Efron B., Tibshirani R.J., An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

[11] Gallais A., Rives M., Detection, number and effects of QTLs for a complex character, Agronomie 13 (1993) 723–738.

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

[13] Hospital F., Charcosset A., Marker-assisted introgression of quantitative trait loci, Genetics 147 (1997) 1469–1485

[14] Hospital F., Moreau L., Lacoudre F., Charcosset A., Gallais A., More on the efficiency of marker-assisted selection, Theor. Appl. Genet. 95 (1997) 1181–1189.

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

[16] Kearsey M.J., Hyne V., QTL analysis: a simple 'marker-regression' approach, Theor. Appl. Genet. 89 (1994) 698–702.

[17] Knapp S.J., Bridges W.C., Using molecular markers to estimate quantitative trait locus parameters: power and genetic variances for unreplicated and replicated progeny, Genetics 126 (1990) 769–777.

[18] Knott S.A., Prediction of the power of detection of marker-quantitative trait locus linkages using analysis of variance, Theor. Appl. Genet. 89 (1994) 318–322.

[19] Lande R., Thompson R., Efficiency of marker assisted selection in the improvement of quantitative traits? Genetics 124 (1990) 743–756.

[20] Lander E.S., Botstein D., Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps, Genetics 121 (1989) 185–189.

[21] Lebreton C.M., Visscher P.M., Empirical nonparametric bootstrap strategies in quantitative trait loci mapping: conditioning on the genetic model, Genetics 148 (1998) 525–535.

[22] Mangin B., Goffinet B., Comparison of several confidence intervals for QTL location, Heredity 78 (1997) 345–353.

[23] Mangin B., Goffinet B., Rebai A., Constructing confidence intervals for QTL location, Genetics 138 (1994) 1301–1308.

[24] Martinez O., Curnow R.N., Estimating the location and the sizes of the effects of quantitative trait loci using flanking markers, Theor. Appl. Genet. 85 (1992) 480–488.

[25] Melchinger A.E., Utz H.F., Schön C.C., Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects, Genetics 149 (1998) 383–403.

[26] Moreau L., Charcosset A., Hospital F., Gallais A., Marker-assisted selection efficiency in populations of finite size, Genetics 148 (1998) 1353–1365.

[27] Ooijen van J.W., Accuracy of mapping quantitative trait loci in autogamous species, Theor. Appl. Genet. 84 (1992) 803–811.

[28] Paterson A.H., De Verna J.W., Lanini B., Tanksley S.D., Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecific cross of tomato, Genetics 124 (1990) 735–742.

[29] Rebai A., Goffinet B., Mangin B., Comparing power of different methods for QTL detection, Biometrics 51 (1995) 87–99.

[30] Simpson S.P., Correction: detection of linkage between quantitative trait loci and restriction fragment length polymorphism using inbred lines, Theor. Appl. Genet. 85 (1992) 110–111.

[31] Soller M., Genizi A., The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations, Biometrics 34 (1978) 46–55.

[32] Soller M., Genizi A., Brody T., On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in a cross between inbred lines, Theor. Appl. Genet. 47 (1976) 35–39.

[33] Visscher P.M., Haley C.S., Power of a chromosomal test to detect genetic variation using genetic markers, Heredity 81 (1998) 317–326.

[34] Visscher P.M., Thompson R., Haley C.S., Confidence intervals in QTL mapping by bootstrapping, Genetics 143 (1996) 1013–1020.

[35] Visscher P.M., Haley C.S., Thompson R., Marker-assisted introgression in backcross breeding programs, Genetics 144 (1996) 1923–1932

[36] Walling G.A., Visscher P.M., Haley C.S., A comparison of bootstrap methods to construct confidence intervals in QTL mapping, Genet. Res. Camb. 71 (1998) 171–180.

[37] Xu S., Further investigation on the regression method of mapping quantitative trait loci, Heredity 80 (1998) 364–373.