DOI: 10.1534/genetics.103.021683

Modifying the Schwarz Bayesian Information Criterion to Locate Multiple Interacting Quantitative Trait Loci

Małgorzata Bogdan,*,† Jayanta K. Ghosh†,‡ and R. W. Doerge†,§,1

**Institute of Mathematics, Wroclaw University of Technology, 50-370 Wroclaw, Poland,* ‡*Indian Statistical Institute, Calcutta 700035, India,* †*Department of Statistics, Purdue University, West Lafayette, Indiana 47907 and* §*Department of Agronomy, Purdue University, West Lafayette, Indiana 47907*

Manuscript received August 27, 2003
Accepted for publication March 5, 2004

ABSTRACT

The problem of locating multiple interacting quantitative trait loci (QTL) can be addressed as a multiple regression problem, with marker genotypes being the regressor variables. An important and difficult part in fitting such a regression model is the estimation of the QTL number and respective interactions. Among the many model selection criteria that can be used to estimate the number of regressor variables, none are used to estimate the number of interactions. Our simulations demonstrate that epistatic terms appearing in a model without the related main effects cause the standard model selection criteria to have a strong tendency to overestimate the number of interactions, and so the QTL number. With this as our motivation we investigate the behavior of the Schwarz Bayesian information criterion (BIC) by explaining the phenomenon of the overestimation and proposing a novel modification of BIC that allows the detection of main effects and pairwise interactions in a backcross population. Results of an extensive simulation study demonstrate that our modified version of BIC performs very well in practice. Our methodology can be extended to general populations and higher-order interactions.

POPULAR methods for mapping quantitative trait loci (QTL) include interval mapping (Lander and Botstein 1989), composite interval mapping (Zeng 1993, 1994) and multiple QTL mapping (MQM; Jansen 1993; Jansen and Stam 1994). These statistical methods do not allow the location of QTL in situations when there are no main effects for the respective QTL, but there are (epistatic) interactions with other QTL (genes) that influence the quantitative trait. Epistatic QTL are known to play important roles in many disease studies, such as cancer (Fijneman *et al.* 1996, 1998), and it is also suspected that they play a key role in the evolutionary process (Wolf *et al.* 2000).

A direct solution to detecting epistatic QTL is to search for several QTL simultaneously and fit an appropriate multiple regression model with interactions. However, the utility of such an approach, which is referred to as a multidimensional version of interval mapping, called multiple interval mapping (MIM; Kao *et al.* 1999), is limited by two interconnected issues. The first is the requirement of deciding how many terms (main effects and epistasis) should be included in the model. The second issue is the computational complexity of the search over the space of possible multidimensional models. To avoid these problems Jannink and Jansen (2001) and Boer *et al.* (2002) proposed one-dimensional genome searches as a means of mapping epistatic QTL. In particular they proposed an interesting extension of MQM by addressing a crucial problem pertaining to the choice of marker cofactors. By including all available markers in a regression equation and using a Bayesian approach to penalize large values of the corresponding regression coefficients many of the previously mentioned issues are eliminated. The disadvantage of this method is that, when detecting epistatic QTL, it requires the choice of "the effective dimension" (*i.e.*, number of QTL) for epistatic interactions, which has strong influence on the power of detection.

An alternative way to approach the problem of mapping epistatic QTL relies on developing new methods for reducing the numerical complexity of MIM. In recent work Carlborg *et al.* (2000), Nakamichi *et al.* (2001), and Broman and Speed (2002) use random search methods to accelerate the search over the class of possible multidimensional models. The results from their approach hold great potential for further progress in solving the problem of the computational complexity of MIM.

Regardless of which method we use to search the genome for QTL we need to solve the problem of estimating QTL number, which in turn directly affects the dimensionality of the model space. The standard way of deciding how many main and interacting (QTL) effects should appear in the model relies on using many statistical tests (see Kao *et al.* 1999). A disadvantage of this approach is that it allows the comparison of only nested

[1]*Corresponding author:* Department of Statistics, 1399 Math Bldg., Purdue University, West Lafayette, Indiana 47907.
E-mail: doerge@purdue.edu

models. It is also unclear how to adjust the significance thresholds for each consecutive test.

Model selection criteria have been used as an alternative approach for the problem of model selection in QTL mapping. Two easy-to-compute model selection criteria that are often employed in statistics are the Akaike information criterion (AIC; Akaike 1974) or the Schwarz Bayesian information criterion (BIC; Schwarz 1978). These criteria belong to the family of the so-called *penalized maximum likelihood methods* and are based on the recommendation of choosing the model for which the likelihood of the data minus the penalty for the model dimension obtains the maximal value. These criteria were used by Jansen (1993) and Jansen and Stam (1994) to choose marker covariates for MQM and by Piepho and Gauch (2001), Nakamichi *et al.* (2001), Ball (2001), Broman and Speed (2002), and Siegmund (2003) to directly estimate QTL number. For a review and discussion of model selection methods as applied to QTL mapping see Balding *et al.* (2002) or Silanpää and Corander (2002).

Piepho and Gauch (2001) investigated many model selection criteria via simulation. In their study different criteria were used to choose pairs of markers flanking QTL. Their results suggest that out of the considered criteria BIC has the best properties and can be recommended for the estimation of the number of QTL with main effects. Broman and Speed (2002), however, recommend a modification of BIC to select markers strongly associated with the trait. Contrary to Piepho and Gauch (2001) they use BIC to choose single markers instead of pairs. Broman and Speed (2002) observe that in this situation the original BIC has a tendency to overestimate the QTL number. To solve the problem of the overfitting they propose a modification of BIC, with a larger penalty for model dimension. Simulations reported in Broman and Speed (2002) show that their modified version of BIC performs very well and detects the correct model more often than composite interval mapping does (Zeng 1993, 1994).

While both of the methods put forth by Piepho and Gauch (2001) and Broman and Speed (2002) can be used to estimate the number of QTL with main effects, they do not generalize directly to the situation where interaction terms appear in the model. Our extensive simulations (Bogdan and Doerge 2003) showed that the phenomenon of overfitting becomes even more significant when we allow interaction terms to appear in the model without the related main effects.

In the present work we concentrate on BIC, which, according to the QTL simulation study of Piepho and Gauch (2001) and our independent simulations, performs better than other popularly used model selection criteria. In particular, we recall the Bayesian roots of BIC and explain the reasons why this criterion, when used to select single markers, has a tendency to overestimate the model dimension. To address this issue we follow the approach suggested by Ball (2001) and propose an easy modification of BIC that relies on taking into account the realistic prior distribution on the set of compared models. In comparison to Ball (2001) we extend the method to cover models with interactions and calibrate the prior to gain the control over the type I error of our procedure. An extensive simulation study verifies that our proposed criterion deals very well with the problem of overfitting the model and allows the detection of main effects and pairwise interactions in a backcross population. While our proposal is based on QTL mapping in a backcross population, our methodology can be extended to general populations and to higher-order interactions.

METHODS

Consider a backcross population where q_{ij} denotes the genotype of the ith individual at the jth QTL: $q_{ij} = -\frac{1}{2}$ if the ith individual is homozygous at the jth QTL and $q_{ij} = \frac{1}{2}$ if it is heterozygous. We assume that the relationship between the trait value Y_i and QTL genotypes is given by a normal regression model,

$$Y_i = \mu + \sum_{j=1}^{m} \beta_j q_{ij} + \sum_{1 \le j < l \le m} \gamma_{jl} q_{ij} q_{il} + \varepsilon_i, \qquad (1)$$

where m is the QTL number and $\varepsilon_i \sim N(0, \sigma^2)$ is the environmental noise. The second summation in our model corresponds to pairwise epistatic interactions. The formulation of the model allows some of the coefficients β_j and γ_{jl} to be zero to accommodate cases when there are QTL that are not involved in epistatic effects. It also addresses the scenario when QTL might not have their own main effects, yet influence the quantitative trait by interacting with other genes, (*i.e.*, epistasis). Later we use p to denote the number of QTL with main effects and q to denote the number of nonzero epistatic terms.

We rely on MIM (Kao *et al.* 1999) to simultaneously locate multiple QTL. This method requires fitting the model (1) for a dense grid of possible QTL positions. For each of the possible genomic locations the genotypes of the putative QTL are inferred using the genotypes of flanking markers and the EM algorithm (Dempster *et al.* 1977) is employed to estimate parameters of the model (1). The locations for which the fitted model yields the largest likelihood are subsequently chosen.

A first step in the reduction of the complexity of MIM sometimes relies on identifying interesting genomic regions on the basis of an initial, relatively coarse search. In the Bayesian setting this approach was suggested by Sen and Churchill (2001), who used an initial scan based on a 10-cM pseudo-marker grid. However, for the situation where an accurate genetic map exists a natural

approach is to base the initial search on the net of marker positions and then use more refined methods (*e.g.*, MIM) to search in the neighborhood of the chosen markers. Ball (2001), Broman and Speed (2002), Yi *et al.* (2003a), and Xu (2003) successfully search over markers to locate multiple QTL and are justified in doing so on the basis of the fact that flanking markers absorb all the information associated with the QTL (Whittaker *et al.* 1996).

If we reduce MIM to a search over markers, then the problem of the QTL location reduces to the problem of choosing the best model of the form

$$Y_i = \mu + \sum_{j\in I}\beta_j X_{ij} + \sum_{(u,v)\in U}\gamma_{uv}X_{iu}X_{iv} + \varepsilon_i, \qquad (2)$$

where X_{ij} denotes the genotype of the ith individual at the jth marker; I is a certain subset of the set $\mathcal{N} = \{1, \ldots, N_m\}$, where N_m is the number of available markers; and U is a certain subset of $\mathcal{N} \times \mathcal{N}$. For a backcross population the random variables $X_{iu}X_{iv}$ correspond to the epistatic terms that are not correlated to any of the main effects. In particular, $X_{iu}X_{iv}$ is not correlated to either X_{iu} or X_{iv} even if the uth and vth markers are statistically dependent via linkage. Thus, the epistatic effects are statistically not confounded with any of the main effects, and in most cases they will be detected only if the epistatic interactions are present.

One difficulty in fitting model (2) is the estimation of the number of main effects and interaction terms to be included in the model. There is a vast statistical literature on the choice of the number of terms in a linear model (see Miller 1990 or McQuarrie and Tsai 1998) and there are many model selection criteria that can be used for this purpose. As mentioned earlier Broman and Speed (2002) and Piepho and Gauch (2001) recommend using the Schwarz BIC (Schwarz 1978) to estimate the number of QTL with main effects. In a general statistical context BIC recommends choosing the model that maximizes the expression

$$S = \log L(Y|\theta) - \frac{1}{2}k \log n, \qquad (3)$$

where θ is the vector of model parameters, $L(Y|\theta)$ is the likelihood of the data, k is the number of parameters (dimension of θ), and n is the sample size. BIC belongs to the wide class of the so-called penalized maximum-likelihood methods and the second term in this criterion, $\frac{1}{2}k \log n$, is called the penalty for the complexity of the model. An important advantage of BIC is that for a wide range of statistical problems, and in particular for multiple regression, it is consistent (*i.e.*, when the sample size grows to infinity, the probability of choosing the right model converges to 1). In the context of linear regression, maximizing S is equivalent to minimizing

$$\mathrm{BIC} = n \log\left(\frac{\mathrm{RSS}}{n}\right) + k \log n, \qquad (4)$$

where RSS is the residual sum of squares from regression.

Rationale for modifying BIC: Broman and Speed (2002) report that the original BIC, when used to select single markers with significant main effects, has a tendency to overestimate QTL number. On the basis of work not shown here (Bogdan and Doerge 2003) we have found that the tendency to overestimate QTL number becomes more significant when the portion (or entirety) of the genome under investigation increases. To understand this further we compare the rates at which the number of different models increases as the number of available markers increases. Our rationale is based on the observation that the number of possible models of the particular form (2), involving k distinct markers, is equal to $\binom{N_m}{k}$, where N_m is the total number of available markers. Thus, when k is much smaller than N_m, the number of models involving k markers increases with N_m approximately like N_m^k. The difference in the numbers of possible "small" and "large" models increases quickly with N_m, and for large N_m the probability of choosing models with many components, just by random chance, is relatively high. Furthermore, for a large number of interaction terms, Bogdan and Doerge (2003) show that the original BIC has a tendency to choose models with epistatic terms even when in reality there is no epistasis.

The phenomenon of overestimation itself suggests the way the standard model selection criteria should be modified to make them useful for QTL mapping. Namely, the high rate at which the number of multidimensional models increases, when the number of available markers increases, suggests that the penalty for the model dimension should increase with this number. This condition is satisfied, for example, by criteria proposed by Broman and Speed (2002) and Siegmund (2003). Second, the fact that there are many more interaction terms than the main effects suggests that the penalty for including an interaction should be larger than the penalty for including a main effect. Following these two suggestions we modify BIC by supplementing it with a realistic prior distribution on the set of possible models. Taking advantage of the fact that BIC is the approximation to the Bayesian rule for the choice of the "best" model we denote by $\theta_i = (\mu, \beta_1, \ldots, \beta_{p(i)}, \gamma_1, \ldots, \gamma_{q(i)}, \sigma)$ the vector of parameters of the ith linear model, M_i, given by Equation 2. Here $p(i)$ and $q(i)$ denote the number of main effects and interaction terms involved in M_i. We assign a certain prior distribution for θ_i and denote the density of this distribution by $f(\theta_i)$. Moreover, let us denote the prior probability of the ith model by $\pi(i)$. Given that $L(Y|\theta_i, M_i)$ denotes the likelihood of the data given the vector of parameters θ_i, let $p(Y|M_i)$ denote the likelihood of the data given the model M_i,

$$p(Y|M_i) = \int L(Y|\theta_i, M_i) f(\theta_i|M_i)\, d\theta_i. \quad (5)$$

The posterior probability of the ith model, given the data, is

$$P(M_i|Y) = \frac{\pi(i)p(Y|M_i)}{\sum_{j=1}^{l}\pi(j)p(Y|M_j)}, \quad (6)$$

where l is the number of possible models.

The Bayesian rule recommends choosing the model for which the posterior probability $P(M_i|Y)$ is the largest (see Schwarz 1978). Since the denominator in Equation 6 is the same for all considered models, Bayes' rule recommends choosing the model for which $\pi(i)p(Y|M_i)$ is the largest. The BIC criterion neglects the prior probabilities $\pi(i)$ of different models and approximates $\log p(Y|M_i)$ by $\log L(Y|\hat{\theta}_i, M_i) - \frac{1}{2}(p(i) + q(i) + 2)\log n$, where $\hat{\theta}_i$ is the maximum-likelihood estimator of θ_i, and $p(i) + q(i) + 2$ is the number of estimated parameters [*i.e.*, $p(i) + q(i)$ for main and epistatic effects, and 2 for μ and σ]. Neglecting $\pi(i)$ corresponds to assigning the same prior probability to all considered models. While in many applications this approach is well justified, in the context of QTL mapping it lends itself to assigning unrealistically high prior probabilities to the events where many regressors are involved [*e.g.*, when 200 markers are available, the number of different models involving 100 main effects is $\binom{200}{100} \approx 9.05 \times 10^{58}$ and the prior probability of the event that 100 regressors are involved is $>10^{56}$ times larger than the prior probability of the event that there is just one regressor]. Motivated to improve on this we suggest supplementing BIC with a more realistic prior distribution, π, on the class of possible models, and choosing the model for which

$$\tilde{S}(i) = \log \pi(i) + \log L(Y|\hat{\theta}_i, M_i) - \frac{1}{2}(p(i) + q(i) + 2)\log n \quad (7)$$

obtains a maximum.

In the context of multiple regression

$$\log L(Y|\hat{\theta}_i, M_i) = -\frac{n}{2}\log \mathrm{RSS}_i + C(n),$$

where $C(n)$ is the constant dependent only on n, and maximizing (7) is equivalent to minimizing the quantity

$$S(i) = n \log \mathrm{RSS}_i + (p(i) + q(i))\log n - 2 \log \pi(i).$$

Prior distribution π: Assume N_m markers are available, and therefore N_m potential regressors and $N_e = (N_m(N_m - 1))/2$ potential interaction terms. The number of all models of the form (2) that can be constructed using subsets of N_m markers is equal to $2^{N_m+N_e}$. To assign prior probabilities to these models we follow the standard solution proposed in George and McCulloch (1993). Namely, we assign the probability α to the event that the ith main effect appears in the model and probability ν to the event that the jth interaction term appears in the model. Our prior distribution assumes that particular terms enter the model independently of others and for a particular model M_i involving $p(i)$ main effects and $q(i)$ interactions we obtain

$$\pi(M_i) = \alpha^{p(i)}\nu^{q(i)}(1 - \alpha)^{N_m - p(i)}(1 - \nu)^{N_e - q(i)}.$$

This choice of prior implies that the prior distributions on the number of main effects and epistatic terms are binomial with parameters N_m and α, and N_e and ν, respectively.

For simplicity we consider α and ν as $\alpha = 1/l$, $\nu = 1/u$, where l and u are certain natural numbers, and restate the prior distribution as

$$\log \pi(M_i) = C(N_m, N_e, l, u) - p(i)\log(l - 1) - q(i)\log(u - 1),$$

where $C(N_m, N_e, l, u)$ is a constant dependent on N_m, N_e, l, and u. Incorporating this prior distribution into the BIC [modified Schwarz BIC (mBIC)] allows the following rule: choose the model that minimizes

$$\mathrm{mBIC}(i) = n \log \mathrm{RSS}_i + (p(i) + q(i))\log n + 2p(i) \times \log(l - 1) + 2q(i)\log(u - 1). \quad (8)$$

The expected values of the prior distribution for the number of main effects are equal to N_m/l and N_e/u for the number of interaction terms. Therefore, since the choice of l and u should reflect our prior knowledge on the QTL number, the values of l and u should be relatively small when we expect many QTL and large when we expect only few. Extensive simulations were performed for the purpose of investigating the standard values of l and u when we have no prior knowledge on the QTL number. We let l and u take on values in such a way that for the sample sizes $n \geq 200$ the probability of type I error (detecting at least one QTL when there are none) does not exceed 5%. We observed that when markers are densely spaced (distance between markers is not >20 cM) we can obtain our aim by keeping the expected values of the number of main effects and interaction terms at a constant level close to 2. In particular, and as is seen next, in our simulations we used values $l \approx N_m/2.2$ and $u \approx N_e/2.2$. In the appendix we present results of some theoretical calculations that support our empirical choice of l and u. These calculations yield approximate bounds on the type I error of our procedure and demonstrate that the proposed choice of l and u solves the problem of multiple comparisons and allows control of the type I error. In comparison to the original BIC the penalty in our proposed/modified criterion involves additional terms $2p(i)\log((N_m/2.2) - 1)$ and $2q(i)\log((N_e/2.2) - 1)$. A similar additional penalty appears in the criterion proposed by Siegmund (2003), who approaches the problem of QTL mapping differently by treating it as a change-point problem. These

TABLE 1

Simulation models

Model	p	Main effects [chromosome, position (cM), β]	q	σ^2	h^2
1	0	—	0	1	0
2	1	(1, 5, 1)	0	1	0.2
3	0	—	1	1	0.195
4	2	(1, 24, 1.5), (1, 56, 1.25)	0	1	0.59
5	2	(1, 24, 1.5), (1, 56, −1.25)	0	1	0.31
6	2	(1, 20, 1.5), (1, 50, 1.25)	0	1	0.59
7	2	(1, 20, 1.5), (1, 50, −1.25)	0	1	0.3
8	2	(1, 20, 1.5), (1, 60, 1.25)	0	1	0.58
9	1	(1, 5, 1)	1	1	0.33
10	0	—	3	1	0.55
11	5	(1, 71, 1.5), (2, 49, 1.25), (3, 27, 1), (4, 8, 0.75), (5, 31, 0.5)	0	1	0.58
12	7	(1, 20, 0.76), (1, 60, 0.76), (2, 20, 0.76), (2, 60, −0.76) (3, 40, 0.76), (4, 20, 0.76), (5, 0, 0.76)	0	1	0.5
13	12	(c, 55, 0.5) for $c = 1, \ldots, 12$	0	1	0.43
14	12	(c, 55, 0.5) for $c = 1, \ldots, 12$	0	0.3	0.71
15	12	(c, 55, 0.5) for $c = 1, \ldots, 12$	0	0.09	0.89
16	2	(1, 71, 1.5), (2, 49, 1)	5	1	0.63
17	5	(1, 24, 1), (1, 96, 1), (2, 5, 1), (3, 5, 0.75), (4, 5, 0.5)	2	1	0.58

The number of QTL with main effects β_j is denoted by p, and q is the number of epistatic terms with effects γ_{jl}, as defined in model (1). The environmental noise is denoted $\varepsilon_i \sim N(0, \sigma^2)$. Broad sense heritability is h^2, and the epistatic effects are as described in Table 2.

additional terms make our criterion similar to the risk inflation criterion (RIC) proposed by Foster and George (1994) in which the penalty for including k orthogonal regressors is equal to $2k \log t$, where t is the total number of available regressors. Note, however, that when n tends to infinity these additional terms are overshadowed by the BIC penalty $(p(i) + q(i)) \log n$ and, contrary to RIC, our criterion has the asymptotic properties of the BIC (*i.e.*, consistency).

SIMULATIONS

We employ computer simulations to evaluate the applicability of our proposed modification to the BIC criterion. Marker and QTL genotypes are simulated for a backcross population using 12 chromosomes of the length 100 cM for sample sizes $n = 200$ and $n = 500$. The number of QTL with main effects ranges between 0 and 12, and the number of epistatic terms between 0 and 5 (Tables 1 and 2). Models 4, 5, and 11–14 (Table 1) are included to allow for a direct comparison to the results of Broman and Speed (2002), as indicated by model 12, and to the results of Piepho and Gauch (2001; models 4, 5, 11, 13, and 14). Since we are interested in how our proposed criterion adjusts to the number of available markers, we search for QTL over 1, 5, and 12 chromosomes and use marker spacings of 5, 10, and 20 cM. The number of available markers and interaction terms, as well as the corresponding values of l and u for each of these experiments, is specified in Table 3. The forward selection procedure (see, *e.g.*, Miller 1990) is used to search the space of possible multidimensional models. At each consecutive step we test all terms (main and interaction) not yet in the model and choose the one whose presence in the model yields the lowest value of the modified BIC criterion (Equation 8; mBIC). To save computational time the procedure is stopped after 30 steps and the resulting 31 models are evaluated on the basis of minimizing the mBIC (Equation 8). The number of steps is restricted to 30 since the largest model we use in the simulations has only 12 terms. Actually, we observe that for all the cases that we considered, the mBIC criterion was minimized by models with <20 terms and that increasing the number of steps above 20 had no influence on the results. However, in real data studies, when one does not want to bound the QTL number, we suggest using a larger number of steps.

RESULTS

The results of searching over 1, 5, and 12 100-cM chromosomes, respectively, with markers spaced every 10 cM are shown in Tables 4–6, while Table 7 reports the results for varying marker distances. The number of correctly identified terms (corr. id.), averaged across 100 simulations, and the average number of false positives (extr.) are reported. The false positives that occur are divided into categories depending on their linkage to true QTL. Following Piepho and Gauch (2001) we classify the main effect to be correct if it corresponds to a marker lying within 15 cM of the true QTL. If

TABLE 2

Details of epistatic effects employed in simulation (Table 1)

Model	q	Epistatic effects (QTL1; QTL2; γ)
3	1	(1, 5; 1, 90; 2)
9	1	(2, 5; 3, 5; 2)
10	3	(1, 71; 2, 49; 3), (3, 27; 4, 8; 2.5), (5, 31; 6, 35; 2)
16	5	(3, 27; 4, 8; 2.5), (5, 31; 6, 35; 2), (7, 5; 8, 5; 1.5), (9, 5; 10, 5; 1), (11, 5; 12, 5; 0.75)
17	2	(5, 5; 6, 5; 2), (7, 5; 8, 5; 1)

QTLi (i = 1, 2) denotes the position of the ith QTL (chromosome and QTL location). The number of epistatic terms and their effects are denoted by q and γ, respectively, and are as described in model (1).

two markers from the neighborhood of one QTL are chosen, one of these markers is arbitrarily classified as extraneous. Epistatic terms are classified as correct if both markers involved lie within 15 cM of the true QTL. For the no-QTL model (1) the percentage of replicates for which the model with no QTL was chosen is reported. While the 15-cM margin is somewhat arbitrary it accommodates our situation well and illustrates the performance of our criterion. Recall that our main goal is the estimation of QTL number and not the precise location of QTL. If we use a narrower range (*i.e.*, <15 cM), then some of the properly identified terms will be classified as extraneous due to the relatively large error of localization of weak QTL that is inherent to all QTL mapping procedures.

Our modification to BIC performs very well (Tables 4–7) in practice, adjusts appropriately to the number of available markers under consideration, and rarely overestimates. Furthermore, in all of the examples we considered the probability of incorrectly detecting at least one QTL, when there are none, does not exceed 0.06 and the average number of extraneous QTL, which are not linked to true QTL, rarely exceeds 0.10. We also observe that the average number of extraneous epistatic terms never exceeded 0.05. This confirms our expectations that in the backcross population epistatic effects are usually detected only when they really exist. Since we set the expected values of the prior distribution for the number of main effects and interaction terms to be equal to 2.2, our criterion more easily identifies models with a small number of terms. The properties of our proposed criterion quickly improve with increasing sample size. Therefore, the accuracy of detecting small models increases (see models 1, 6, and 7 in Table 4) as does the ability to correctly identify models with larger numbers of QTL (see models 12, 13, 16, and 17 in Table 4). We are aware that the chance of correctly identifying QTL depends on its heritability. In other words, when the variance of the error is equal to 1.0 and the sample size is $n = 200$, our criterion usually detects main effects with coefficients $\beta \geq 0.76$ (the heritability of the single QTL with such a β is 0.13) and interaction terms with $\gamma \geq 2$ (broad sense heritability of 0.20 with just one such epistatic term in the model) even when they appear in larger models. When the sample size is increased to $n = 500$ our criterion usually detects main effects with $\beta \geq 0.50$ (individual $h^2 \geq 0.06$) and interaction terms with $\gamma \geq 1.5$ (individual $h^2 \sim 0.12$). The proposed criterion (mBIC) works particularly well if QTL are located close to markers (compare models 4 and 6, and 5 and 7, in Tables 4–6 and models 4 and 8 in Table 7). When QTL are located in the middle of an interval defined by two markers it is sometimes the case that both flanking markers are chosen, which partially explains the relatively large number of *false* positives for models 4 and 15. An additional reason for the sometimes larger number of extraneous linked QTL is a statistical error of localization of weak QTL. In some cases the correct model was appropriately identified, but the chosen markers were slightly farther apart from the true QTL than our set limit of 15 cM. On the basis of this reasoning some of the *false* positives correspond to correctly identi-

TABLE 3

Penalty coefficients l and u used in the modified BIC (mBIC)

No. of chromosomes	Marker spacing (cM)	N_m	$N_e = \frac{N_m(N_m - 1)}{2}$	l	u
1	10	11	55	5	25
5	10	55	1,485	25	675
12	10	132	8,646	60	3,930
12	5	252	31,626	115	14,375
12	20	72	2,556	33	1,162

N_m denotes the number of markers and N_e denotes the number of available interactions.

TABLE 4

Results from 100 simulations that each search over 12 100-cM chromosomes with markers spaced every 10 cM

		Main terms				Epistatic terms				
Model	n	p	Corr. id.	Extr. linked	Extr. unlinked	q	Corr. id.	Extr. both linked	Extr. one linked	Extr.both unlinked
1	200	0	0.95	—	0.03	0	—	—	—	0.02
1	500	0	0.99	—	0.01	0	—	—	—	0.00
2	200	1	1.00	0.01	0.02	0	0.00	0.00	0	0.02
3	200	0	—	0.00	0.01	1	0.95	0.01	0.00	0.01
4	200	2	1.97	0.31	0.02	0	—	0.00	0.00	0.04
5	200	2	1.98	0.06	0.02	0	—	0.00	0.01	0.03
6	200	2	2.00	0.10	0.02	0	—	0.00	0.00	0.04
6	500	2	2.00	0.02	0.02	0	—	0.00	0.00	0.01
7	200	2	2.00	0.07	0.01	0	—	0.00	0.00	0.03
7	500	2	2.00	0.01	0.01	0	—	0.00	0.00	0.02
9	200	1	1.00	0.00	0.03	1	0.92	0.03	0.00	0.01
10	200	0	—	0.01	0.01	3	2.86	0.03	0.03	0.01
11	200	5	4.08	0.18	0.00	0	—	0.00	0.02	0.00
12	200	7	5.02	0.23	0.02	0	—	0.01	0.04	0.01
12	500	7	6.99	0.13	0.01	0	—	0.00	0.03	0.00
13	200	12	2.39	0.31	—	0	—	0.02	—	—
13	500	12	9.68	0.47	—	0	—	0.02	—	—
14	200	12	9.53	0.75	—	0	—	0.02	—	—
15	200	12	11.9	0.63	—	0	—	0.04	—	—
16	200	2	1.95	0.03	0.01	5	2.08	0.12	—	—
16	500	2	2.00	0.01	0.02	5	3.46	0.07	—	—
17	200	5	3.67	0.18	0	2	0.80	0.04	0.00	0.01
17	500	5	4.80	0.18	0	2	1.32	0.03	0.00	0.00

p is the true number of main effects, q is the true number of epistatic terms, n is the sample size, Corr. id. denotes the average number of correctly identified terms, Extr. linked denotes the average number of extraneous terms that are linked to true QTL, and Extr. unlinked denotes the average number of extraneous terms that are not linked to true QTL.

fied, but incorrectly localized QTL. Comparing results of our simulations with the results reported in Piepho and Gauch (2001) and Broman and Speed (2002) we observe, for models with only main effects, that our modification of BIC (mBIC) performs similarly to the criteria proposed in these earlier articles. More importantly, however, our criterion allows the detection of epistatic terms whereas the criteria of Piepho and Gauch (2001) and Broman and Speed (2002) do not.

DISCUSSION

The method proposed in this article can be viewed as a simplification of standard Bayesian methods used for QTL mapping. In a series of articles Satagopan and Yandell (1996), Satagopan *et al.* (1996), Heath (1997), Uimari and Hoeschele (1997), Silanpää and Arjas (1998), Stephens and Fisch (1998), and Yi and Xu (2000) use the full Bayesian approach and Markov chain Monte Carlo simulations to estimate posterior distributions of QTL locations and other parameters in the regression model. Yi *et al.* (2003a), Xu (2003), and Kilpikari and Silanpää (2003) reduce the number of parameters generated by Markov chain Monte Carlo (MCMC) by restricting the search to marker positions. Yi and Xu (2002) and Yi *et al.* (2003b) extend the standard Bayesian MCMC approach to search for epistatic QTL. The common feature shared by the works of these authors is that they require multiple generations from the conditional distributions of all parameters in the regression model and are very computationally demanding. Moreover, as noted by Ball (2001), "a major challenge remains to obtain a rapidly converging sampler for the full Bayesian model." Sen and Churchill (2001) avoided using MCMC by employing an independent sample Monte Carlo approach to generate multiple versions of pseudo-marker genotypes on the dense grid of genomic locations. They computed weights for each pseudo-marker realization by integrating out parameters of the related regression models and then used them to approximate the posterior distribution of the QTL locations. Our method, similar to the methods of Ball (2001) and Broman and Speed (2002), is a further simplification of Bayesian methodology and seems to be particularly useful when one needs to search over a large space of possible models with interactions.

TABLE 5

Results from 100 simulations that each search over one 100-cM chromosome with markers spaced every 10 cM

		Main effects			Epistatic terms		
Model	n	p	Corr. id.	Extr.	q	Corr. id.	Extr.
1	200	0	0.96	0.03	0	—	0.03
1	500	0	0.94	0.04	0	—	0.02
2	200	1	0.99	0.04	0	—	0.02
3	200	0	—	0.02	1	0.99	0.02
4	200	2	2.00	0.59	0	—	0.03
5	200	2	2.00	0.26	0	—	0.05
6	200	2	2.00	0.13	0	—	0.02
6	500	2	2.00	0.07	0	—	0.02
7	200	2	2.00	0.09	0	—	0.01
7	500	2	2.00	0.05	0	—	0.02

p is the true number of main effects, q is the true number of epistatic terms, n is the sample size, Corr. id. denotes the average number of correctly identified terms, and Extr. denotes the average number of extraneous terms.

The modified BIC that is presented here is closer than the original BIC to the concept of Bayesian thinking since it introduces the prior distribution on the number of main effects and epistatic terms. We concentrate mainly on the situation when there are no specific expectations on the number of QTL and calibrate the prior so as to gain control over the type I error of our procedure. However, we strongly suggest that in the case when some prior information is available it should be included and the penalty should be adjusted accordingly. To estimate the type I error in that case one could use computer simulations or the permutation method of Churchill and Doerge (1994).

In principle, the modified version of BIC suggested in this article could be used to approximate posterior probabilities of different models according to the formula

$$P(M_i|Y) \approx \frac{\exp(-\text{mBIC}(i)/2)}{\sum_{j=1}^{l}\exp(-\text{mBIC}(j)/2)}, \qquad (9)$$

where l is the number of possible models (see also Ball 2001). While we are very much aware of the importance of this formulation, which could allow one to estimate the uncertainty related to the choice of the best model and to use Bayesian averaging to estimate main and epistatic effects, we point out that due to the huge number of possible models with interactions it is practically impossible to compute its denominator. To reduce the number of terms in the Equation 9 one could apply Occam's window algorithm proposed by Raftery *et al.* (1997), which relies on discarding models that receive little support from the data. However, the corresponding search procedure proposed in Madigan and Raftery (1994) seems to be inadequate in our setting due to the large number of nonnested models. In practice one may reduce the number of models considered by performing a separate search for each pair of chromosomes, which in turn is usually good enough to detect pairwise interactions. But even in this case, the number of possible models with interactions will usually be too large to apply Equation 9.

To solve the problem of multiplicity of models and to identify the best one, we applied forward selection procedure, which is simple and quick. Our simulations, as well as results reported in Broman and Speed (2002), show that forward selection performs very well in this setting. We are, however, aware that there are some

TABLE 6

Results from 100 simulations that each search over five 100-cM chromosomes with markers spaced every 10 cM

	Main effects				Epistatic terms				
Model	p	Corr. id.	Extr. linked	Extr. unlinked	q	Corr. id.	Extr. both linked	Extr. one linked	Extr. both unlinked
1	0	0.96	—	0.02	0	—	—	—	0.02
2	1	0.99	0.02	0.00	0	—	0.00	0.00	0.02
3	0	—	0.00	0.01	1	0.96	0.01	0.01	0.00
4	2	1.97	0.37	0.03	0	—	0.00	0.00	0.01
5	2	1.97	0.08	0.02	0	—	0.00	0.00	0.01
6	2	2.00	0.14	0.02	0	—	0.00	0.00	0.01
7	2	2.00	0.07	0.02	0	—	0.00	0.00	0.01
9	1	0.99	0.03	0.01	1	0.93	0.04	0.01	0.00
11	5	4.27	0.25	0.00	0	—	0.04	—	—
12	7	5.55	0.27	—	0	—	0.03	—	—

Sample size $n = 200$, p is the true number of main effects, q is the true number of epistatic terms, Corr. id. denotes the average number of correctly identified terms, Extr. linked denotes the average number of extraneous terms that are linked to true QTL, and Extr. unlinked denotes the average number of extraneous terms that are not linked to true QTL.

TABLE 7

Results of the search over 12 100-cM chromosomes based on 100 simulations and the sample size $n = 200$

		Main effects				Epistatic terms				
Model	D (cM)	p	Corr. id.	Extr. linked	Extr. unlinked	q	Corr. id.	Extr. both linked	Extr. one linked	Extr. both unlinked
1	5	0	0.98	—	0.01	0	—	—	—	0.01
1	20	0	0.95	—	0.03	0	—	—	—	0.02
2	5	1	0.99	0.00	0.00	0	—	0.00	0.00	0.01
2	20	1	1.00	0.02	0.02	0	—	0.00	0.00	0.03
4	5	2	2.00	0.11	0.01	0	—	0.00	0.01	0.03
4	20	2	1.87	0.54	0.03	0	—	0.00	0.03	0.01
8	20	2	2.00	0.03	0.04	0	—	0.00	0.07	0.02
10	5	0	—	0.01	0.03	3	2.94	0.03	0	0.01
10	20	0	—	0.01	0.02	3	2.21	0.10	0.02	0.02
12	5	7	4.66	0.19	0.01	0	—	0.00	0.00	0.00
12	20	7	5.23	0.43	0.08	0	—	0.01	0.03	0.04

D is the marker spacing, p is the true number of main effects, q is the true number of epistatic terms, Corr. id. denotes the average number of correctly identified terms, Extr. linked denotes the average number of extraneous terms that are linked to true QTL, and Extr. unlinked denotes the average number of extraneous terms that are not linked to true QTL.

particular cases (and a real analysis is always a particular case) when the forward selection procedure does not detect the best model. Thus, although statistically we do not expect much improvement by replacing forward selection with a more refined search strategy, we still recognize the need for further research in this direction.

Although this article is concerned solely with detecting main effects and pairwise interactions, theoretically the proposed method can be directly generalized to identify higher-order interactions. To retain control over the type I error of the corresponding procedure, it is anticipated that higher-order interactions should be penalized even more than pairwise epistatic terms. However, the utility of this approach needs to be verified by additional research, since there are two main difficulties related to any extensions of our work. First is the numerical complexity of the search over a rapidly increasing number of models with higher-order interactions, which can most likely be addressed by developing a suitable search strategy and increasing computer power. The second issue is more difficult and of a more theoretical nature. If we do not have prior expectations on the number of main and epistatic effects the method outlined in this article can be used to control the overall type I error. In this case, when we increase the potential number of regressors by including higher-order interactions, we must also increase the penalties for main effects and pairwise interactions. Thus, an attempt to detect higher-order interactions will result in decreasing power of detection of simpler effects and can be offset only by larger sample sizes. When some prior information on the number of main effects and interactions is available the power will be less affected since the method can be used in a subjective way via an appropriate adjustment of the penalties.

In this article we did not address the problem of missing marker data. Currently in the QTL mapping literature three methods exist, which are designed to solve this problem by using genotypes of neighboring markers. They include Haley and Knott (1992) regression, the E-M algorithm of Jansen and Stam (1994), or multiple imputations of missing genotypes proposed by Sen and Churchill (2001) and Ball (2001). We believe that the application of any of these methods will leave the mBIC unaffected by a moderate proportion of missing marker data. The missing data methods can also be used to apply mBIC to search for QTL within intermarker intervals.

The method proposed in this article selects markers strongly associated with the trait and does not explicitly use the information from the distance between them. Therefore, in principle the mBIC approach is not sensitive to map errors. However, the application of any of the missing data methods will make our method sensitive to map errors in the same way as standard interval mapping. Our method can be also influenced by selective genotyping and genotyping errors, since selective genotyping will change the correlation structure in the design matrix and might result in partial confounding of epistatic and main effects. However, our approach is able to select the proper markers out of many strongly correlated neighbors; therefore we believe that it is also robust to any partial confounding of main and epistatic effects. The influence of genotyping errors will depend on the marker information that is affected. In our mBIC criterion, as well as in other standard model selection criteria, the information on the data appears only in RSS. Thus, we do not expect a significant difference between our criterion and others with respect to the sensitivity to genotyping errors.

Locating QTL, and more importantly their interactions, remains an open problem in both the QTL mapping and statistical communities. Current multiple interval mapping methods are plagued by two intimately related issues. First is the problem of estimating the number of QTL and their interactions. And second is the related issue of searching over the space of all possible multidimensional models that comprise the computationally complex space. Realizing that the second issue is impacted by the approach of the first issue, we have presented a model selection criterion that allows more accurate assessment than the original BIC criterion, from which we started. Using simulations in a backcross setting we demonstrate that the mBIC does well to locate multiple interacting QTL. Extensions of our method to more general designs are possible and are currently under investigation. Furthermore, the proposed criterion can be used outside the context of genetics to estimate the number of additive effects and interactions in the general framework of multiple regression.

We thank Jim Berger, Andreas Futschik, and Joanna Szyda for helpful discussions and two anonymous reviewers for comments that greatly improved the presentation of the results. We also thank Adam Zagdanski and Monika Horobiowska-Kaczmarz for help in performing simulations. The results reported in this article were presented at the Seventh Purdue Symposium on Statistics, June 19–24, 2003. R.W.D. is funded in part by grants from the U.S. Department of Agriculture-Initiative for Future Food and Agricultural Systems (00-52100-9615) and the National Science Foundation (0115109-MCB).

LITERATURE CITED

Akaike, H., 1974 A new look at the statistical model identification. IEEE Trans. Automat. Control **AC-19:** 716–723.

Balding, D., A. D. Carothers, Y. L. Marchini, L. R. Cardon, A. Vetta *et al.*, 2002 Discussion on the meeting on statistical modelling and analysis of genetic data. J. R. Stat. Soc. B **64:** 737–775.

Ball, R., 2001 Bayesian methods for quantitative trait loci mapping based on model selection: approximate analysis using the Bayesian information criterion. Genetics **159:** 1351–1364.

Boer, M. P., C. J. F. ter Braak and R. C. Jansen, 2002 A penalized likelihood method for mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **162:** 951–960.

Bogdan, M., and R. W. Doerge, 2003 Mapping multiple interacting quantitative trait loci with multidimensional genome searches. Technical Report 04-03. Department of Statistics, Purdue University, West Lafayette, IN.

Broman, K. W., and T. P. Speed, 2002 A model selection approach for the identification of quantitative trait loci in experimental crosses. J. R. Stat. Soc. B **64:** 641–656.

Carlborg, Ö., L. Andersson and B. Kinghorn, 2000 The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics **155:** 2003–2010.

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 EM algorithm. J. R. Stat. Soc. B **39:** 1–38.

Fijneman, R. J. A., S. S. de Vries, R. C. Jansen and P. Demant, 1996 Complex interactions of new quantitative trait loci, *Sluc1*, *Sluc2*, *Sluc3*, and *Sluc4*, that influence the susceptibility to lung cancer in the mouse. Nat. Genet. **14:** 465–467.

Fijneman, R. J. A., R. C. Jansen, M. A. Van der Valk and P. Demant, 1998 High frequency of interactions between lung cancer susceptibility genes in the mouse: mapping of Sluc5 to Sluc14. Cancer Res. **58:** 4794–4798.

Foster, D. P., and E. I. George, 1994 The risk inflation criterion for multiple regression. Ann. Stat. **22:** 1947–1975.

George, E. I., and R. E. McCulloch, 1993 Variable selection via Gibbs sampling. J. Am. Stat. Assoc. **88:** 881–889.

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.

Heath, S. C., 1997 Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. Am. J. Hum. Genet. **61:** 748–760.

Jannink, J.-L., and R. Jansen, 2001 Mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **157:** 445–454.

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.

Kao, C.-H., Z-B. Zeng and R. D. Teasdale, 1999 Multiple interval mapping for quantitative trait loci. Genetics **152:** 1203–1216.

Kilpikari, R., and M. J. Silanpää, 2003 Bayesian analysis of multilocus association in quantitative and qualitative traits. Genet. Epidemiol. **25:** 122–135.

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

Madigan, D., and A. E. Raftery, 1994 Model selection and accounting for model uncertainty in graphical models using Occam's window. J. Am. Stat. Assoc. **89:** 1535–1546.

McQuarrie, A. D. R., and C.-L. Tsai, 1998 *Regression and Time Series Model Selection.* World Scientific Publishers, Singapore.

Miller, A. J., 1990 *Subset Selection in Regression.* Chapman & Hall, London.

Nakamichi, R., Y. Ukai and H. Kishino, 2001 Detection of closely linked multiple quantitative trait loci using genetic algorithm. Genetics **158:** 463–475.

Piepho, H.-P., and H. G. Gauch, Jr., 2001 Marker pair selection for mapping quantitative trait loci. Genetics **157:** 433–444.

Raftery, A. E., D. Madigan and J. A. Hoeting, 1997 Bayesian model averaging for linear regression models. J. Am. Stat. Assoc. **92:** 179–191.

Satagopan, J. M., and B. S. Yandell, 1996 Estimating the number of quantitative trait loci via Bayesian model determination. Special Contributed Paper Session on Genetic Analysis of Quantitative Traits and Complex Diseases, Biometric Section, Joint Statistical Meetings, Chicago.

Satagopan, J. M., B. S. Yandell, M. A. Newton and T. C. Osborn, 1996 Bayesian model determination for quantitative trait loci. Genetics **144:** 805–816.

Schwarz, G., 1978 Estimating the dimension of a model. Ann. Stat. **6:** 461–464.

Sen, S., and G. A. Churchill, 2001 A statistical framework for quantitative trait mapping. Genetics **159:** 371–387.

Serfling, R., 1980 *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Siegmund, D., 2003 Model selection in irregular problems: applications to mapping QTLs. Biometrika (in press).

Stephens, D. A., and R. D. Fisch, 1998 Bayesian analysis of quantitative trait locus data using reversible jump Markov chain Monte Carlo. Biometrics **54:** 1334–1347.

Silanpää, M. J., and E. Arjas, 1998 Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics **148:** 1373–1388.

Silanpää, M. J., and J. Corander, 2002 Model choice in gene mapping: what and why. Trends Genet. **18:** 301–307.

Uimari, P., and I. Hoeschele, 1997 Mapping linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. Genetics **146:** 735–743.

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.

Wolf, J. B., E. D. Brodie, III and M. J. Wade (Editors), 2000 *Epistasis and the Evolutionary Process.* Oxford University Press, New York.

Xu, S., 2003 Estimating polygenic effects using markers of the entire genome. Genetics **163:** 789–801.

Yi, N., and S. Xu, 2000 Bayesian mapping of quantitative trait loci for complex binary traits. Genetics **155:** 1391–1403.
Yi, N., and S. Xu, 2002 Mapping quantitative trait loci with epistatic effects. Genet. Res. **79:** 185–198.
Yi, N., V. George and D. B. Allison, 2003a Stochastic search variable selection for identifying multiple quantitative trait loci. Genetics **164:** 1129–1138.
Yi, N., S. Xu and D. B. Allison, 2003b Bayesian model choice and search strategies for mapping interacting quantitative trait loci. Genetics **165:** 867–883.
Zeng, Z-B., 1993 Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. Proc. Natl. Acad. Sci. USA **90:** 10972–10976.
Zeng, Z-B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Communicating editor: J. B. Walsh

APPENDIX: BOUND FOR THE TYPE I ERROR

Our procedure recommends choosing the model that maximizes the criterion

$$\tilde{S}(i) = \log L(Y|\hat{\theta}_i) - \frac{1}{2}(p(i) + q(i))\log n - p(i)\log(l-1) - q(i)\log(u-1). \quad \text{(A1)}$$

The number of all possible one-dimensional models [models for which $p(i) + q(i) = 1$] is equal to $N_m + N_e$, where, as before, N_m is the number of available markers and $N_e = (N_m(N_m - 1))/2$ is the number of possible interactions. Let $\tilde{S}_1$ denote the maximum of the criterion (A1) over all such one-dimensional models and let $\tilde{S}_0 = \log L_0(Y|\hat{\mu}, \hat{\sigma})$ be the value of the criterion for the null model involving no markers ($p + q = 0$). Let $D = p + q$ be the number of terms in the model chosen by our procedure. It holds that

$$P(D > 0) = P(\tilde{S}_1 > \tilde{S}_0) + P(D > 1, \tilde{S}_1 \le \tilde{S}_0).$$

We bound the probability of the first, dominating term of the right-hand side of the above equality, under the null hypothesis of no QTL.

Consider a given one-dimensional model M_i and a corresponding value of our criterion

$$\tilde{S}_{M_i} = \log L(Y|\hat{\theta}_i) - \frac{1}{2}\log n - (\log(l-1) \text{ or } \log(u-1)).$$

The model M_i will be preferred over the model with no QTL if $\tilde{S}_{M_i} > \tilde{S}_0$, or equivalently

$$2\log\frac{L(Y|\hat{\theta}_i)}{L_0(Y|\hat{\mu}, \hat{\sigma})} > \log n + 2(\log(l-1) \text{ or } \log(u-1)). \quad \text{(A2)}$$

Under the null hypothesis of no QTL $2\log L(Y|\hat{\theta}_i)/L_0(Y|\hat{\mu}, \hat{\sigma})$ has asymptotically χ^2 distribution with 1 d.f. (Serfling 1980). Thus, $P(\tilde{S}_{M_i} > \tilde{S}_0)$ is asymptotically equivalent to

$$2P\left(Z > \sqrt{\log n + 2(\log(l-1) \text{ or } \log(u-1))}\right),$$

where Z is a $N(0, 1)$ random variable.

Note that $\tilde{S}_1 > \tilde{S}_0$ if Equation A2 holds for at least one of the one-dimensional models. Therefore, by Bonferroni inequality, for any $\varepsilon > 0$ and sufficiently large n it holds

$$P(\tilde{S}_1 > \tilde{S}_0) \le 2N_m P\left(Z > \sqrt{\log n + 2\log(l-1)}\right) + 2N_e P\left(Z > \sqrt{\log n + 2\log(u-1)}\right) + \varepsilon. \quad \text{(A3)}$$

For each $x > 0$ it holds that

$$P(Z > x) \le \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\frac{1}{x}.$$

Thus (A3) yields

$$P(\tilde{S}_1 > \tilde{S}_0) \le \frac{2N_m}{(l-1)\sqrt{2\pi n(\log n + 2\log(l-1))}} + \frac{2N_e}{(u-1)\sqrt{2\pi n(\log n + 2\log(u-1))}} + \varepsilon.$$

For the proposed values $l = N_m/2.2$ and $u = N_e/2.2$ the right-hand side of the above inequality is approximately equal to

$$\frac{4.4}{\sqrt{2\pi n}}\left(\frac{1}{\sqrt{\log n + 2\log(l-1)}} + \frac{1}{\sqrt{\log n + 2\log(u-1)}}\right). \quad \text{(A4)}$$

Using the proposed values of l and u allows one to eliminate N_m and N_e from the bound numerator and thus helps solve the multiple-comparisons problem.

For the sample size $n = 200$ and N_m and N_e used in our experiments the bound given by (A4) takes values from the interval between 0.0574 (for $N_m = 252$ and $N_e = 31{,}626$; 12 100-cM chromosomes with markers spaced every 5 cM) and 0.0801 (for $N_m = 11$ and $N_e = 55$; one 100-cM chromosome with markers spaced every 10 cM), which gives a satisfactory approximation for the empirical type I error obtained from simulations.