

Use of Score Statistics for Model Selection in Mapping Multiple Quantitative Trait Loci

Cecelia Laurie*, Shengchu Wang[†],
Luciana Aparecida Carlini-Garcia[‡] and Zhao-Bang Zeng^{†§}

*Department of Mathematics

University of Alabama, Tuscaloosa, AL 35487-0350, USA

[†] Department of Statistics & Bioinformatics Research Center
North Carolina State University, Raleigh, NC 27695-7566, USA

[‡] Instituto Agronômico de Campinas

Centro de Grãos e Fibras, Campinas, SP, Brazil

[§] Department of Genetics

North Carolina State University, Raleigh, NC 27695-7566, USA

Running head: Use of Score Statistics for Model Selection

Corresponding address: Zhao-Bang Zeng
Bioinformatics Research Center
Department of Statistics
North Carolina State University
Raleigh, NC 27695-7566, U.S.A.

Telephone number: 919-515-1942

Fax number: 919-515-7315

E-mail: zeng@stat.ncsu.edu.

Key words: Score statistic, Multiple Interval Mapping,
Model Selection

ABSTRACT

One of the major issues in mapping multiple quantitative trait loci is that of model selection. Many methods for mapping QTL scan positions across the genome and test for statistical significance to decide whether to admit QTL in the model. A common problem with these methods is the determination of a genome-wide threshold value for the test statistic. Recently a model- and data-based resampling procedure using score statistics was introduced to empirically calculate the relevant threshold. In this paper we extend the score statistic analysis to multiple interval mapping (MIM) procedure and perform a simulation study to assess the performance of mapping multiple QTL using a sequential MIM search with model selection based on the use of the score statistic threshold. Our results show that this procedure provides an effective, consistent genome-wide threshold for admitting a QTL to be included in an MIM model. We discuss the performance of the procedure with respect to model size, false discovery rate, power to detect QTL, confidence in position estimates, and behavior of parameter estimates. An important result is that when we relax the genome-wide threshold from the nominal 5% significance level to 20% for model selection, the statistical power to detect QTL can be significantly increased and at the same time the false discovery rate is only mildly affected and is well controlled. Based on the simulation result, we recommend the use of genome-wide threshold at 20% significance level calculated from the score resampling method for model selection in mapping multiple QTL.

The problem of identifying the genetic factors underlying complex quantitative traits has been a major research focus in quantitative genetics. With the availability of a well-saturated genetic map of molecular markers, it is possible to identify a major part of the structure of the genetic architecture of quantitative traits and to estimate the associated parameters. For standard experimental designs (such as backcross, F_2 , or use of recombinant inbred lines), the inference of genetic architecture of a quantitative trait involves the number, genetic locations, main and interaction effects of QTL (quantitative trait loci).

Many statistical methods for mapping QTL focus predominantly on detecting individual QTL effects and assigning these effects to genomic regions. Both interval mapping (IM) (Lander & Botstein, 1989) and composite interval mapping (CIM) (Zeng, 1994; Jansen & Stam, 1994) detect QTL effects at different genomic regions separately. Although based on a single gene model, these methods have been shown to be successful in identifying a few significant QTL. However, complex traits will involve multiple QTL as well as interactions (epistasis) among the QTL. This dictates that the search for and mapping of multiple epistatic QTL need to be performed at multiple intervals simultaneously. Approaches that model multiple QTL will improve statistical power to identify more minor and complex QTL, better separate linked QTL and allow the examination of interactions between QTL. Several methods for considering multiple-QTL simultaneously have been proposed (e.g. Kao & Zeng, 1997, Kao *et al.*, 1999, and Zeng *et al.*, 1999; Broman & Speed, 2002). Bayesian approaches have also been proposed (Satagopan *et al.*, 1996; Sillanpaa & Arjas, 1998; Sen & Churchill, 2001; Yi *et al.*, 2005).

One of the major issues in multiple-QTL approaches is that of model selection, i.e. the choice of which and how many parameters to include in the model. The choice of good criteria to make the model selection is an open problem. Broman & Speed (2002) focus on the selection of the number and approximate locations of QTL using multiple regression on markers, arguing that this model selection stage is a prerequisite and hence of primary importance (considering the estimation of effects and precise locations of secondary importance). They propose a modified BIC criterion, BIC_δ , for

determining the model where the δ factor increases the penalty on the size of the model over the standard BIC criterion. MIM (Kao, *et al.*, 1999; Zeng, *et al.*, 1999), rather than restricting the analysis to markers, searches the positions of multiple QTL across the genome. MIM can improve the precision of estimating QTL positions, identify epistasis, and provide appropriate and integral estimation of individual QTL effects, variance and covariance contributions. The model selection is based on sequential testing for admitting (or deleting) a QTL in the model.

A common problem in many of the methods for mapping QTL that scan positions across the genome (including MIM) is the determination of a threshold value for the test statistic. Many factors, such as genome size, genetic map density, informativeness of markers, and proportion of missing data, may affect the distribution of the test statistic. Using pointwise significance levels is inadequate since the entire genome is tested for the presence of a QTL, leading to multiple tests that are not independent. For single QTL, theoretical approximations have been developed to determine threshold and power (Lander & Botstein 1989; Dupuis & Siegmund 1999; Rebai *et al.* 1994). Theoretical approximations are based on rigid distributional requirements. Concerns about multiple tests and distributional assumptions can be bypassed by using resampling methodology to find empirical thresholds. Churchill & Doerge (1994) introduced permutation testing. The idea is to replicate the original analysis many times on data sets generated by randomly reshuffling the original trait data while leaving the marker data unchanged. Doerge & Churchill (1996) proposed residual-based and conditional-based extensions of the permutation-based method for estimating threshold values for detecting minor QTL effects while accounting for effects of known major QTL. For MIM, where model selection is involved, Zeng *et al.* (1999) proposed using a residual bootstrap resampling method for hypothesis testing. However, these resampling methods are computationally intensive.

Zou *et al.* (2004) proposed a model- and data-based resampling procedure using score statistics to empirically calculate the relevant threshold. It is less computationally demanding than permutation tests or bootstrap resampling (on the order of 10^2 or higher), more accurate than theoretical approximations, and applicable to more

complicated designs and models. They assessed the performance of their method by investigating type I error and power by simulating the null and alternative models for a single QTL in an F_2 cross. The assessment of their method for multiple QTL included a sequential and also a simultaneous search for two QTL in a backcross.

In this paper we extend the score statistic analysis to MIM and, through a simulation study of backcross population, evaluate the performance of mapping multiple QTL using a sequential MIM search with the model selection based on the use of the score statistic threshold. The fit of model will be evaluated based on a number of measures including false discovery rate. The procedures and analysis also apply to other experimental designs, such as F_2 . Based on this study, we also make general recommendations on practical procedures for QTL analysis based on MIM.

METHODS

The model and its likelihood: Assuming m putative QTL in a backcross population, the MIM model is defined by

$$y_i = \mu + \sum_{k=1}^m a_k x_{ik} + \sum_{k \neq l \in \{1, \dots, m\}} \delta_{kl} \gamma_{kl} x_{ik} x_{il} + \varepsilon_i.$$

Here y_i is the phenotypic trait value for individual i for $i = 1, \dots, n$ while x_{ik} is a coded variable denoting the genotype of putative QTL k using the Cockerham scale:

$$x_{ik} = \begin{cases} \frac{1}{2} & \text{if the genotype of the QTL } k \text{ for the individual } i \text{ is homozygous} \\ -\frac{1}{2} & \text{if the genotype of the QTL } k \text{ for the individual } i \text{ is heterozygous} \end{cases}$$

The variable x_{ik} is unobserved, but its conditional probability can be analyzed given observed marker genotypes and specific positions for the QTL (Jiang & Zeng, 1997). Parameters include the mean (μ), the marginal effects of the putative QTL (a_k 's), the variance (σ^2) of the residual effects ε_i , assumed to be normally distributed with mean zero, and the epistatic effects. The epistatic effect between putative QTL k and l is denoted γ_{kl} . We use only a subset of all QTL pairs where the indicator variable δ_{kl} takes the value one if QTL k and QTL l interact and takes the value zero otherwise.

Assume there are m QTL in the model with t epistatic effects. Let \mathbf{v} denote the vector of positions for the m QTL and let θ denote the vector of parameters (main and epistatic effects, overall mean μ and the variance σ^2 of the residual effects).

Since the genotypes of an individual at the putative QTL are not usually observed (but marker genotypes are), the model contains missing data and thus the likelihood function of the data given the model is a mixture of normal distributions:

$$L(\theta; \mathbf{v}) = \prod_{i=1}^n \left[\sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right]$$

The term in the square brackets is the likelihood for individual i . Here $\phi(y_i | \mu_j, \sigma^2)$ denotes a normal density function for y_i with mean μ_j and variance σ^2 . The probability density of each individual is a mixture of 2^m possible normal density functions with different means, μ_j , one for each of the 2^m possible multiple-QTL genotypes. To infer the distribution of QTL genotypes, we assume there is no crossover interference and also that double recombination events within an interval between markers are very rare and can be ignored. The mixing proportion p_{ij} is the probability of each multilocus genotype conditioned on marker data (Jiang & Zeng, 1997). The mean μ_j is the genotypic value of the j^{th} multilocus QTL genotype. These genotypic values can be found by $\mu_j = \mu + \mathbf{D}_j \mathbf{E}$ where \mathbf{E} is the column vector of effects parameters (a_k 's and γ_{kl} 's) and \mathbf{D}_j is the j^{th} row of the genetic design matrix \mathbf{D} specifying the configuration of x 's, the coded genetic variables associated with each a and γ for the j^{th} QTL multilocus genotype. (See Appendix for more detailed descriptions of \mathbf{D} and \mathbf{E} .)

We denote the log-likelihood of individual i by $l_i(\theta; \mathbf{v})$ and the overall log-likelihood by $l(\theta; \mathbf{v})$. Hence we have

$$l_i(\theta; \mathbf{v}) = \ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\}$$

$$l(\theta; \mathbf{v}) = \sum_{i=1}^n \ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\}$$

When m is large, the number of possible mixture components (QTL genotypes) can become very large. Many of the probabilities (p_{ij} 's) can be very small. In a practical implementation of MIM, we adopt a selection procedure to choose a subset of 'significant' p_{ij} components for evaluation and normalize so that the sum of the probabilities equals 1. (See Zeng, *et al.* 1999.)

Parameter estimation: In parameter estimation, the finite normal mixture model can be treated as an incomplete-data problem by regarding the trait and markers as observed data and the QTL as missing data. Kao & Zeng (1997) and Zeng *et al.* (1999) describe a procedure to obtain maximum likelihood parameter estimates using an expectation/maximization (EM) algorithm (as well as estimates of the asymptotic variance-covariance matrix).

For complex models with many QTL and parameters, scanning every position in the marker intervals leads to many maximum likelihood calculations. This becomes intractable without an efficient algorithm for finding the maximum likelihood estimators. The EM algorithm, although numerically stable, can become infeasible due to very slow convergence. To speed convergence but retain stability, Rai and Matthews (1993) introduced a hybrid iterative method combining generalized EM and Newton-Raphson method (NR) (see also McLachlan & Krishnan 1997). After a few steps of EM, the results are used as initial values in a modified Newton-Raphson algorithm using derivatives of the conditional expected complete-data log-likelihood rather than the complete log-likelihood. As long as the step-size in the modified NR stages is chosen to satisfy certain conditions, the algorithm ensures that the likelihood increases at each step. X. Qin & Z.-B. Zeng (unpublished) implemented this algorithm, which they called GEM-NR, in the multi-trait, multi-environment setting. The GEM-NR algorithm was shown to have faster convergence and to have a tendency to converge at higher values of the likelihood. Details concerning the GEM-NR algorithm and its advantages will be discussed in a separate paper. We have implemented this algorithm for the one trait–one environment model and have found it to have somewhat faster convergence than the EM algorithm. However, there does not seem to be a big advantage over the EM algorithm for low to moderate complexity of the model. The Appendix includes a brief description of how the algorithm is used for our purposes.

Model search and model selection: Our focus in this paper will be on analyzing the use of score statistics for identifying main QTL effects. We perform a forward sequential search, adding one QTL at a time to the model. At each stage, a grid of positions across the genome are tested. (The grid uses 1 or 2 cM intervals, avoiding

regions within 5 cM on either side of already identified QTL in the model.) For each position in the grid, the MLE's for the parameters are computed assuming that there is a QTL at that position and including already identified QTL in the model. A statistic is computed for testing the null hypothesis that the effect parameter for the QTL at the new position is zero. The putative QTL position with the maximum test statistic is added to the model *if the statistic exceeds a specified threshold*. This process is repeated until the maximum test statistic does not exceed its corresponding threshold.

The key ingredient in the model selection is choosing an appropriate test statistic and corresponding threshold value. Zou, *et al.* (2004) proposed using score statistics and a resampling procedure for determining the threshold.

Assume that there are $m-1$ QTL already identified. Let d represent a position being tested for the m^{th} QTL. Consider the model with m QTL parameters $\theta = (\eta, \beta) = (\theta_1, \dots, \theta_{m-1}, \beta, \mu, \sigma^2)$. Here β is the effect parameter for the 'new' m^{th} QTL and $\eta = (\theta_1, \dots, \theta_{m-1}, \mu, \sigma^2)$ is considered the vector of nuisance parameters.

The procedure for adding the m^{th} QTL in the model consists of the following steps. More complete explanations and justifications of the procedure can be found in Zou, *et al.* (2004).

1. *Compute the MLE's* for the m effects parameters as well as for μ and σ^2 as discussed in the Parameter Estimation section.
2. *Compute the Score Statistic* for the position.
3. *Compute the threshold* based on a large number, say N , of resampled score statistics.
 - (a) Compute the $100(1-\alpha)^{\text{th}}$ percentile of the set of N resampled score statistics to determine the threshold value.
 - (b) Accept the position being tested as identifying a new QTL if the score statistic for the position exceeds the threshold value.

Step 2: Compute the Score Statistic

Let $U(d)$ denote the score function for β evaluated at $\beta = 0$ and $\tilde{\eta}$ where $\tilde{\eta}$ is the MLE of η under the null hypothesis. It can be shown that $n^{-\frac{1}{2}}U(d)$ has the same asymptotic distribution as $n^{-\frac{1}{2}}\sum_{i=1}^n U_i(d)$ where $U_i(d)$ involves only the information from the i^{th} subject and hence the $U_i(d)$ ($i = 1, \dots, n$) are independent zero-mean random variables for any given d . The fact that the score function can be approximated by a sum of independent random vectors will allow us to derive its large-sample distribution. The approximations would be found by replacing the unknown parameters by their sample estimators under the null hypothesis, i.e. evaluating at $(\beta, \eta) = (0, \tilde{\eta})$. We thus have:

$$\hat{U}_i(d) = U_{\beta,i}(0, \tilde{\eta}; d) - \left(\frac{\partial^2 l(0, \tilde{\eta}; d)}{\partial \beta \partial \eta} \right) \left(\frac{\partial^2 l(0, \tilde{\eta}; d)}{\partial \eta^2} \right)^{-1} U_{\eta,i}(0, \tilde{\eta}; d)$$

where $\tilde{\eta}$ denotes the maximum likelihood estimator of η for the restricted model $\beta = 0$.

$$U_{\beta,i}(\beta, \eta; d) = \frac{\partial l_i(\beta, \eta; d)}{\partial \beta}$$

$$U_{\eta,i}(\beta, \eta; d) = \frac{\partial l_i(\beta, \eta; d)}{\partial \eta} = \left(\frac{\partial l_i}{\partial \theta_1}, \dots, \frac{\partial l_i}{\partial \theta_{m-1}}, \frac{\partial l_i}{\partial \mu}, \frac{\partial l_i}{\partial \sigma^2} \right)'$$

$$\hat{U}(d) = \sum_{i=1}^n \hat{U}_i(d)$$

The score statistic for $H_0: \beta = 0$ against $H_1: \beta \neq 0$ at location d is

$$W(d) = \hat{U}'(d) \hat{V}^{-1}(d) \hat{U}(d)$$

where $\hat{V}(d) = \sum_{i=1}^n \hat{U}_i(d) \hat{U}_i'(d)$. The expression for \hat{V} consistently estimates the variance of $n^{-\frac{1}{2}}\hat{U}(d)$. (Here the notation M' refers to the transpose of the matrix M . Note however that, for the backcross model, $\hat{U}(d)$ is a scalar.)

Formulas for the needed first and second partial derivatives are given in the Appendix.

Zou, *et al.* (2004) recommended replacing the restricted MLE with the unrestricted MLE in the evaluation of the score statistic. We have found that the score statistic using the restricted MLE better approximates the behavior of the likelihood ratio statistic

than the score statistic using the unrestricted MLE. (The score statistic and the likelihood ratio statistic are asymptotically equivalent.)

Step 3: Compute the threshold

1. Generate independent samples G_i , $i = 1, 2, \dots, n$ from $N(0, 1)$. Here n is the sample size.
2. Calculate $U^*(d) = \sum_{i=1}^n \widehat{U}_i(d)G_i$, $W^*(d) = U^{*\prime}(d)\widehat{V}^{-1}U^*(d)$, and $S^* = \max_d W^*(d)$.
3. Repeat step 1 (choosing a new sample $\{G_i\}$) and step 2 a large number of times, say N times. Let S_k^* for $k = 1, \dots, N$ denote the k^{th} resampled maximum score statistic.
4. Compute the $100(1 - \alpha)^{th}$ percentile of $\{S_k^* : k = 1, \dots, N\}$ to determine the threshold value.
5. Accept the position being tested as identifying a new QTL if the observed score statistic for the position exceeds the threshold value.

When computing $W^*(d)$, we regard $\widehat{V}(d)$ and the $\widehat{U}_i(d)$ in $U^*(d)$ as fixed and the G_i in $U^*(d)$ as random. In this way, $U^*(d)$ is normal with mean zero and the same expected covariances as those of $\widehat{U}(d)$. Thus $n^{-\frac{1}{2}}U^*(d)$ converges to the same limiting distribution as $n^{-\frac{1}{2}}\widehat{U}(d)$. Consequently the distribution of $W(d)$ can be approximated by the distribution of $W^*(d)$.

Note that the $\widehat{U}_i(d)$ and $\widehat{V}(d)$ used in the resampling calculations are based on the original data and are evaluated once and used repeatedly in step 2; only the G_i 's are changed in each resample. Since it does not involve refitting the model in each iteration, the proposed method is computationally more efficient than the permutation method.

The final model we use in the simulation study is the model from forward selection of QTL, followed by optimization of position estimates for the QTL based on the LOD profile. To optimize the estimates of the QTL positions, the estimates of the positions are updated in turn for each region. For an identified QTL Q_k in the model, the

region between its two neighbor identified QTL or until the ends of the chromosome is scanned to find the position that maximizes the likelihood (conditioned on the current estimates of positions of other QTL in the model). This refinement process is repeated sequentially for each QTL position until there is no change in the estimates of the QTL positions.

Measures of model fit: We investigate several measures of model fit over varying threshold significance levels and LOD support intervals: model size, false discovery rate, power, measures of confidence in position estimates using LOD support intervals, and measures of confidence in parameter estimates.

After a final model is determined, for each simulated QTL we pair it with the nearest QTL on the same chromosome identified in the model. In many simulation studies, false discovery rates and power computations are based simply on pairing identified and simulated QTL. However, this gives no measure of precision; e.g. is the identified QTL “close enough” to the simulated QTL to truly represent the simulated QTL being detected? Some simulation studies consider this by checking if the paired identified QTL is within a certain distance of the simulated QTL. We have chosen to use a criterion based on LOD-support intervals. The reasoning behind this is two-fold: first, the precision of detection depends on the effect size so it is not clear that a fixed distance captures the desirable information; second, there is reason to believe that, even though LOD-support intervals also depend upon effect size, they display relatively stable coverage as confidence intervals. Specifically, Manichaikul, *et al.* (2006) show that, for interval mapping with one QTL, LOD-1.5 support intervals give appropriate coverage as 95% confidence intervals.

A LOD- z support interval around an identified QTL is the longest contiguous interval of positions in which a LOD score is within z of the maximum LOD. (Here the LOD score is computed conditional on all of the other identified QTL being in the model.) In the case of more than one identified QTL on a chromosome or if the QTL position is near the end of the chromosome, the LOD support interval may have an endpoint determined by the position of a neighboring identified QTL or by the end of a chromosome even if the LOD score has not dropped the specified amount.

Let R be the total number of replicates and let M be the number of simulated QTL. (For our cases, $R = 1000$ and $M = 8$.) Let Q_k , $k = 1, \dots, M$ represent the simulated QTL. We will investigate measures of fit based on LOD- z support intervals for $z = 1, 1.5, 2$. For a given k and z , let $\mathfrak{R}_{k,z}$ represent the set of replicates for which the simulated QTL Q_k is paired with an identified QTL and Q_k lies in the LOD- z support interval of the identified QTL. Let $|\mathfrak{R}_{k,z}|$ represent the cardinality of the set $\mathfrak{R}_{k,z}$.

Model size: We compute the percentage of replicates in which a given number of QTL are identified in the model.

False Discovery Rate: An identified QTL is said to be ***misidentified*** if it is either not paired with a simulated QTL or, if it is paired with a simulated QTL, the simulated QTL does not lie in the LOD-support interval of the identified QTL.

$$\text{FDR per replicate} = F_r(z) = \frac{\# \text{ of misidentified QTL in replicate } r}{\text{total } \# \text{ of QTL identified in replicate } r}$$

$$FDR(z) = \frac{1}{R} \sum_{r=1}^R F_r(z)$$

Note that this definition of FDR is based on individual QTL identification and not on correct model size identification.

Power to Detect QTL: For each k and z , compute

$$P(Q_k, z) = \frac{|\mathfrak{R}_{k,z}|}{R}$$

LOD-Support Interval Coverage: For each k and z , compute the percentage

$$C(Q_k, z) = \frac{|\mathfrak{R}_{k,z}|}{\# \text{ of replicates for which } Q_k \text{ is paired with an identified QTL}} \times 100$$

Position Estimates: For each k and z , compute

$$Pos(Q_k, z) = \text{average over replicates in } \mathfrak{R}_{k,z} \text{ of the position estimate for } Q_k$$

$$\mathcal{W}(k, z) = \text{average over replicates in } \mathfrak{R}_{k,z} \text{ of width of LOD-}z \text{ support interval}$$

Bias of position estimates will be assessed. If we interpret the LOD-support interval as a confidence interval, the average width can be interpreted as a measure of the precision of the position estimate.

Parameter Estimates: Given a final model for a replicate, we compute the parameter estimates and the observed Fisher information matrix (the negative of the matrix of second partial derivatives of the log-likelihood evaluated at the MLE's for the parameters). For a simulated QTL, any replicate for which the simulated QTL is not paired with an identified QTL is deleted from consideration. For this subset of replicates associated with the simulated QTL, we compute the mean of the effect size parameter estimates, the observed standard deviation from this mean, and the average of the estimated standard deviations obtained from the inverse of the observed information matrix.

Since parameter estimates are obtained only when the QTL is detected through the model selection process, we expect the effect parameter estimates to be inflated (Beavis 1998, Xu 2003, and Bogdan and Doerge 2005, as well as references in these papers). We investigate the observed bias in the parameter estimates. Since the joint distribution of the maximum likelihood estimates of the parameters tends asymptotically to a multivariate normal distribution, the inverse of the observed Fisher information matrix can be used as the estimated variance-covariance matrix of the parameter estimates. We compare the estimated standard deviations with the observed standard deviations.

Simulations: We simulated two cases of backcross population with a reasonable number of QTL and good genome size: 8 QTL on 9 chromosomes where each chromosome was 110 cM with 12 markers placed every 10 cM. For each case the sample size was 300 individuals and the heritability $h^2 = 0.4$; these values were chosen to roughly mimic practical situations of many reported QTL studies. One thousand (1000) replicates were generated and, for each replicate, 1000 resampling steps were used in computing thresholds. The trait values were generated with a residual error following the standard normal distribution $N(0, 1)$. The cases differed in the distribution of the positions of the QTL and in the additive effect sizes.

Case 1 is “ideal” in that each QTL has the same effect size and there is no linkage among QTL; the results from this case are used as a base for comparison. Case 2 describes a more practical situation where effect sizes vary and there is some linkage (coupling linkage between two QTL on one chromosome and repulsion linkage between

two QTL on another chromosome). The effect sizes for Case 2 were chosen to give a good range of potential power to detect QTL. In addition, to provide more insight into the behavior of the score threshold for differing genome sizes, we simulated a case (Case 3) with 8 QTL, varying effect sizes matching those of Case 2, with a smaller genome size (3 chromosomes). This case also necessarily included more linkage and can thus provide some information on the effect of linkage for the score procedure. Chromosome length, marker spacing, heritability, and sample size were the same as for the other cases. Table 1 provides a description of the position of each QTL (chromosome and position on chromosome) along with the effect size for each of the cases.

RESULTS

Baseline behavior of score statistic and threshold: Interval mapping for detecting one QTL using the likelihood ratio statistic and permutation threshold is well established in the literature. Some initial simulations show that the behavior of the score statistic and associated score threshold is comparable to the behavior of the likelihood ratio statistic with associated permutation threshold. These results reinforce and extend the comparison results of Zou *et al.* (2004). In a simulation of the null model, the score procedure behaves as expected with respect to Type 1 error, with the observed Type I error matching the nominal threshold significance level α . In simulations of one QTL on one chromosome, the power to detect one simulated QTL using the score threshold was comparable to (and slightly higher than) the power using the permutation threshold. Figure 1 provides further evidence of comparable behavior of the score and permutation procedures when searching for the first QTL to add to the model in a multiple QTL situation. Figure 1A shows that the score threshold and permutation threshold for any given threshold significance level α are comparable. Figure 1B shows that the likelihood ratio and score statistic profiles for a genome scan for one QTL are comparable. (The results displayed are for Case 1 but the results for Cases 2 and 3 are similar.)

Consistency of score thresholds in model selection: One of the major issues in multiple-QTL analysis is finding good threshold criteria for model selection. The

score statistic procedure provides a theoretical rationale (as well as a computationally feasible method) for determining data-based genome-wide thresholds for a step-wise search for multiple QTL. A key result is that the score statistic procedure provides a consistent genome-wide threshold for admitting a new QTL in the model. Figure 2 shows that successive score thresholds for a given significance level α are nearly constant (Case 1, equal effects). The thresholds for Case 2 (varying effects, same genome size) exhibit the same consistent behavior as QTL are added to the model with the same genome-wide threshold levels as Case 1 (data not shown). Case 3 also exhibits consistent behavior of the thresholds as QTL are added to the model (with different threshold levels than the other cases since the genome size differs). These results show that the score statistic genome-wide threshold mainly depends on genome size and not on the distribution of effect size or linkage.

Model size: Figure 3 displays the percentage of replicates for which the final model includes a given number of QTL for various significance levels; part A displays the results for Case 1 (equal effects), part B displays the results for Case 2 (varying effects, some linkage), and part C displays the results for Case 3 (smaller genome size, more linkage). For Case 1, there are well defined peaks for models with the correct number (8) of QTL. Significance levels of 15 - 20% have sharper peaks with minimal skew toward larger models; it can be seen that lower significance levels are too conservative. For Case 2, the distributions for most of the significance levels peak at models with 7 QTL in the model. Since the 8th QTL effect is very low, a model with 7 QTL is arguably the "correct" model size. The results for this case emphasize that significance levels below 15% are too conservative and appear to clarify that a significance level around 20% would be a good choice. For Case 3, the effects of linkage can be seen to reduce the power to detect the correct number of QTL. However, the distributions of model sizes for this smaller genome size exhibit patterns similar to the results for the larger genome size and support the basic conclusions that significance levels below 15% are too conservative and that a significance level of around 20% provides a good balance between power to detect QTL and false positives.

False Discovery Rate: Table 2 provides a summary of the FDR for various

significance levels (α) and LOD-support intervals for Case 1, 2 and 3. Note that these are observed FDR rates resulting from using a genome-wide threshold based on score statistics and not on any FDR controlling criteria. These QTL-based false discovery rates in general are considerably lower than the threshold significance level. Comparison of FDR for Case 1, 2 and 3 indicates that the distribution of QTL effect sizes does not have a major effect on FDR, and linkage can increase FDR only slightly. More importantly, FDR is not very sensitive to the increase in significance level (the difference across significance levels from 5% to 30% is about 0.03). The FDR results also indicate that concerns about errors propagating as QTL are added to the model are unfounded when using an appropriate genome-wide threshold.

Power to detect QTL: The fact that the distribution of FDR is rather flat implies that we might be able to increase power substantially without undesirable significant increases in FDR. Table 3 displays the power results for Case 2 using LOD-1.5 support intervals. Power results show that, for medium (QTL 1, 4, 5) and low (QTL 7, 8) effect sizes, increases in power across significance levels can be substantial, up to 10 to 20%, whereas FDR increases only by 2 to 3%. This is an important result. It strongly suggests that we should use more relaxed threshold, such as the threshold at genome-wide 20% significance level, for model selection. The change in power across different LOD- z intervals is only 2 to 5% (data not shown).

Confidence in position estimates: Results show that position estimates from the score statistic procedure are unbiased as expected (Table 4). More importantly, they show that LOD support intervals can reasonably be used as confidence intervals for position estimates.

LOD coverages for a given LOD- z interval across QTL are comparable for all but the lowest effect sizes. For example, for 20% significance level and LOD-1.5 support interval, LOD coverage is 95% or higher for all but the two QTL with lower effect size and lower power for detection (Table 3 and 4). LOD support interval widths (as well as the standard deviations of position estimates) are smaller for larger effects, indicating greater precision of position estimates for QTL with larger signals, as expected.

The effect of differing LOD levels on interval coverage and width can be seen in

Table 4. This table displays position estimate results for 20% significance level but the patterns described above are reflected at all significance levels. Difference in LOD coverage across significance levels is minimal (with a total difference of about 3%). Interval widths do not change much across significance levels for large to medium effects; the difference in widths for a given QTL and LOD- z are within 5 cM (with most within 1 cM) except for the the two lower effect sizes where the widths could differ by 10 cM or more.

Based on these results, we recommend the use of LOD-1.5 intervals as confidence intervals for position estimates for a good balance of FDR, power, coverage, and precision.

QTL effect estimates: Average effect estimates are inflated, with larger bias for smaller effect sizes, as expected due to the model selection process (Beavis 1998, Xu 2003, and Bogdan and Doerge 2005). Results in Table 5 are for significance level 20%. In general, effect estimates are somewhat less biased for more liberal threshold significance levels. (However, for a given QTL, the largest difference between effect estimates across significance levels is less than half of any of the observed standard deviations; data not shown).

Reasons for the expected inflation of parameter estimates due to the model selection procedure have been discussed by several authors (see, for example, references mentioned above as well as references in those papers). Since parameter estimates are obtained only when the QTL is detected through the model selection process, the estimates are based on a censored sample from the asymptotic normal distribution of the maximum likelihood estimates of the parameters. Since we know the power to detect each QTL based on simulation results from the 1000 repetitions, we used statistical theory of censored samples (Cohen 1991) to estimate the expected bias due to censoring. Although this expected bias was comparable to the observed bias for some QTL, including the low effect QTL 7 and 8 (data not shown), we conclude that there are other factors affecting bias in a multi-QTL model selection process, such as linkage and the effect of model size, that are not easily teased apart.

The estimated standard deviations from the inverted observed Fisher information

matrix are consistent across effect sizes. This is to be expected. If the QTL's were closely linked to markers, the standard deviations of the MLE's of the parameters would depend only on the residual error variance, the sample size, and the variance/covariance structure of the genotype indicator variables. The estimated standard deviations in general underestimate the observed standard deviations but are roughly the same order of magnitude. The differences between the observed and estimated can be attributed to the model search process. Chen (2005) performed simulations simultaneously estimating positions and effects for multiple QTL (no genome search) and showed good agreement between observed standard deviations and standard deviations estimated from the inverted observed Fisher information matrix.

DISCUSSION AND RECOMMENDATIONS

One of the major issues in multiple-QTL mapping is that of model selection, i.e. the choice of which and how many parameters to include in the model. The choice of good threshold criteria to make the model selection is an open problem. The score statistic resampling procedure, introduced by Zou *et al.* (2004) and used in conjunction with MIM (Kao, *et al.*, 1999; Zeng, *et al.*, 1999), provides a theoretical rationale (as well as a computationally feasible method) for determining data-based thresholds for a step-wise search for multiple QTL. Like permutation thresholds, the score procedure for determining thresholds is based on the data and, unlike permutation thresholds, is easily extended to mapping multiple QTL. The score threshold appropriately accounts for genome size in the search process via genome scanning in contrast to BIC thresholds (where penalties depend only on sample size).

Our results show that the score statistic procedure provides an effective, consistent genome-wide threshold for admitting a QTL to be included in an MIM model for mapping multiple QTL. Errors do not propagate as QTL are added to the model and the method can effectively control false discovery rates. Appropriate balance between false discovery and power to detect can be adjusted by using different values of the genome-wide significance level parameter α used in determining the threshold. We recommend using α around 20% for giving a good balance between false discovery and power to detect (for each individual QTL as well as for appropriate model size);

results from differing genome sizes (9 chromosomes and 3 chromosomes) support this recommendation. Results further show that LOD support intervals can reasonably be interpreted as confidence intervals for position estimates and thus that the width of such intervals can provide estimates of the precision of QTL location. We recommend the use of LOD-1.5 intervals which give 90 - 95% coverage, except in the case of very small QTL effects.

The fact that the score thresholds are nearly constant as QTL are added to the model can be exploited to save computational time and effort. One could compute the threshold to admit the first QTL and then use it for admitting succeeding QTL without needing further resampling. In this way, one could also afford the computational time to use more resamplings in computing the threshold.

In this study, we have not discussed appropriate and justifiable ways to study epistasis of QTL. This will be pursued in a separate study and published elsewhere. The method developed in this study has been implemented in Windows QTL Cartographer under multiple interval mapping. The score resampling method is interfaced in the software as the default criterion for model selection.

This study is partially supported by the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service, grant number 2005-00754.

LITERATURE CITED

- BEAVIS, W.D., 1998 QTL analyses: power, precision, and accuracy, pp. 145-162 in *Molecular Dissection of Complex Traits*, edited by A.H. PATTERSON. CRC Press, New York.
- BOGDAN, M. and R.W. DOERGE, 2005 Biased estimators of quantitative trait locus heritability and location in interval mapping. *Heredity* **95**: 476-484.
- BROMAN, K. and T. SPEED, 2002 A model selection approach for the identification of quantitative trait loci in experimental crosses. *J.R. Stat. Soc. Ser. B* **64**: 641-656.
- CHEN, Z., 2005 The full EM algorithm for the MLEs of QTL effects and positions and their estimated variances in multiple-interval mapping. *Biometrics* **61**: 474-480.
- CHURCHILL, G.A. and R.W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963-971.
- COHEN, A.C., 1991 *Truncated and Censored Samples*. Marcel Dekker, New York.
- DOERGE, R.W. and G.A. CHURCHILL, 1996 Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**: 285-294.
- DUPUIS, J. and D. SIEGMUND, 1999 Statistical methods for mapping quantitative trait loci from a dense set of markers. *Genetics* **151**: 373-386.
- JANSEN, R.C., 1993 Interval mapping of multiple quantitative trait loci. *Genetics* **135**: 205-211.
- JANSEN, R.C. and P. STAM, 1994 High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**: 1447-1455.
- JIANG, C and Z.-B. ZENG, 1997 Mapping quantitative trait loci with dominant and missing markers. *Genetica* **101**: 47-58.
- KAO, C.-H. and Z.-B. ZENG, 1997 General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when

- using the EM algorithm. *Biometrics* **53**: 653-665.
- KAO, C.-H., Z.-B. ZENG, and R.D. TEASDALE, 1999 Multiple interval mapping for quantitative trait loci. *Genetics* **152**: 1203-1216.
- LANDER, E.S. and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.
- MANICHAIKUL, A., J. DUPUIS, S. SEN, and K.W. BROMAN, 2006 Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* **174**: 481-489.
- McLACHLAN, G.J. and T. KRISHNAN, 1997 *The EM Algorithm and Extensions*. John Wiley, New York.
- RAI, S.N. and D.E. MATTHEWS, 1993 Improving the EM algorithm. *Biometrics* **49**: 587-591.
- REBAI, A., B. GOFFINET, and B. MANGIN, 1994 Approximate threshold of interval mapping tests for QTL detection. *Genetics* **138**: 235-240.
- SATAGOPAN, J.M., B.S. YANDELL, M.A. NEWTON and T.C. OSBORN, 1996 A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. *Genetics* **144**: 805-816.
- SEN, S. and G.A. CHURCHILL, 2001 A statistical framework for quantitative trait mapping. *Genetics* **159**: 371-387.
- SILLANPAA, M.J. and E. ARJAS, 1998 Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* **148**: 1373-1388.
- YI, N., B.S. YANDELL, CHURCHILL, G.A., D.B. ALLISON, E.J. EISEN and D. POMP, 2005 Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. *Genetics* **170**: 1333-1344.
- XU, S., 2003 Theoretical basis of the Beavis effect. *Genetics* **165**: 2259-2268.
- ZENG, Z.B., 1994 Precision mapping of quantitative trait loci. *Genetics* **136**: 1457-

1468.

ZENG, Z.-B., C.-H. KAO and C.J. BASTEN, 1999 Estimating the genetic architecture of quantitative traits. *Genet. Res., Camb.* **74**: 279-289.

ZOU, F., J.P. FINE, J. HU and D.Y. LIN, 2004 An efficient resampling method for assessing genome-wide statistical significance in mapping quantitative trait loci. *Genetics* **168**: 2307-2316.

APPENDIX

The Model and its Likelihood

For the backcross model as described in the paper, assume that there are m QTL in the model with t epistatic effects for a total of $c = m + t$ effects parameters. The likelihood is a mixture of 2^m normal distributions, one for each of the possible missing QTL genotypes.

The individual and the overall log-likelihoods are given by

$$l_i(\boldsymbol{\theta}; \mathbf{v}) = \ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\}$$

$$l(\boldsymbol{\theta}; \mathbf{v}) = \sum_{i=1}^n \ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\}$$

We have that \mathbf{v} denotes the vector of positions of the QTL and $\boldsymbol{\theta}$ denotes the vector of main and epistatic effects parameters along with the overall mean μ and the variance σ^2 of the residual effects. Recall that $\phi(y_i | \mu_j, \sigma^2)$ denotes the normal distribution with mean μ_j and variance σ^2 and that the mixing proportion p_{ij} is the probability of the j^{th} multilocus genotype conditioned on marker data.

Genetic Design Matrix D:

Let $\mathbf{D} = [d_{jp}]_{2^m \times c}$ where row j corresponds to the j^{th} multilocus QTL genotype and column p corresponds to the p^{th} effects parameter. For $p \leq m$, $d_{jp} = \frac{1}{2}$ or $-\frac{1}{2}$ depending upon whether the p^{th} locus in the multilocus QTL genotype j is homozygous or, respectively, heterozygous. For $p > m$, $d_{jp} = \frac{1}{4}$ if the two interacting QTL in the multilocus genotype j are both homozygous or both heterozygous and $d_{jp} = -\frac{1}{4}$ if one of the QTL is homozygous and the other heterozygous.

Effects Matrix E:

Let \mathbf{E} be the $c \times 1$ vector of the effects parameters, i.e. $\mathbf{E} = [a_1; \dots; a_m; \gamma_1; \dots; \gamma_t]$ where a_1, \dots, a_m are the main effects parameters and $\gamma_1, \dots, \gamma_t$ are the epistasis effects parameters.

Genotypic Value μ_j :

Assuming multilocus genotype j , we have that $\mu_j = \mu + \mathbf{D}_j \mathbf{E}$ where \mathbf{D}_j denotes the j^{th} row of \mathbf{D}

Mixing Probabilities Matrix $\mathbf{\Pi}$:

Let π_{ij} denote the probability of multilocus QTL genotype j conditional on marker genotypes and also phenotypic value, for individual i . We have that

$$\pi_{ij} = \frac{p_{ij} \phi(y_i | \mu_j, \sigma^2)}{\sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2)}$$

Let $\mathbf{\Pi}$ be the $n \times 2^m$ matrix whose (i, j) entry is π_{ij} .

For further derivation and discussion of the above quantities, see Kao, *et al.* (1999).

Other Matrices and Notation:

Let \mathbf{T} be the $n \times 2^m$ matrix such that $t_{ij} = y_i - \mu_j$.

Let \mathbf{S} be the $n \times 2^m$ matrix such that $s_{ij} = \frac{t_{ij}^2}{2\sigma^2} - \frac{1}{2}$.

Throughout what follows, the matrix operator $\#$ denotes the Hadamard product of two matrices, i.e. element-wise multiplication. For a matrix G , G_i refers to the i^{th} row of G , $G^{[q]}$ refers to the q^{th} column of G , and G' refers to the transpose of G .

First and second partial derivatives of the log-likelihood functions

Let $\boldsymbol{\theta} = (\theta_1, \dots, \theta_c, \mu, \sigma^2)$ denote the vector of parameters for the model and let \mathbf{v} denote the vector of positions of the QTL. In what follows we will shorten the notation $l(\boldsymbol{\theta}; \mathbf{v})$ and $l_i(\boldsymbol{\theta}; \mathbf{v})$ to $l(\boldsymbol{\theta})$ and $l_i(\boldsymbol{\theta})$ respectively, since it is understood that the likelihood depends on the positions but our focus is on derivatives with respect to the parameters in $\boldsymbol{\theta}$.

Basic building blocks: Assume that $k = 1, \dots, c$.

$$\frac{\partial(y_i - \mu_j)}{\partial \theta_k} = -d_{jk}$$

$$\frac{\partial(y_i - \mu_j)}{\partial \mu} = -1$$

$$\frac{\partial}{\partial \theta_k} \phi(y_i | \mu_j, \sigma^2) = \frac{\partial}{\partial \theta_k} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(y_i - \mu_j)^2}{\sigma^2}} \right] = \phi(y_i | \mu_j, \sigma^2) \frac{1}{\sigma^2} (y_i - \mu_j) d_{jk}$$

$$\frac{\partial}{\partial \mu} \phi(y_i | \mu_j, \sigma^2) = \frac{\partial}{\partial \mu} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(y_i - \mu_j)^2}{\sigma^2}} \right] = \phi(y_i | \mu_j, \sigma^2) \frac{1}{\sigma^2} (y_i - \mu_j)$$

$$\frac{\partial}{\partial \sigma^2} \phi(y_i | \mu_j, \sigma^2) = \frac{\partial}{\partial \sigma^2} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(y_i - \mu_j)^2}{\sigma^2}} \right] = \frac{1}{\sigma^2} \phi(y_i | \mu_j, \sigma^2) \left(\frac{(y_i - \mu_j)^2}{2\sigma^2} - \frac{1}{2} \right)$$

First partial derivatives of individual log-likelihoods

$$\frac{\partial l_i(\boldsymbol{\theta})}{\partial \theta_k} = \frac{\partial}{\partial \theta_k} \left[\ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\} \right] = \sum_{j=1}^{2^m} \pi_{ij} \frac{1}{\sigma^2} (y_i - \mu_j) d_{jk}$$

$$\frac{\partial l_i(\boldsymbol{\theta})}{\partial \mu} = \frac{\partial}{\partial \mu} \left[\ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\} \right] = \sum_{j=1}^{2^m} \pi_{ij} \frac{1}{\sigma^2} (y_i - \mu_j)$$

$$\frac{\partial l_i(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{\partial}{\partial \sigma^2} \left[\ln \left\{ \sum_{j=1}^{2^m} p_{ij} \phi(y_i | \mu_j, \sigma^2) \right\} \right] = \sum_{j=1}^{2^m} \pi_{ij} \frac{1}{\sigma^2} \left(\frac{(y_i - \mu_j)^2}{2\sigma^2} - \frac{1}{2} \right)$$

Second partial derivatives of the full log-likelihood:

We will need the partial derivatives of π_{ij} . In the expressions below we will use the notation $t_{ij} = y_i - \mu_j$ and $s_{ij} = \frac{(y_i - \mu_j)^2}{2\sigma^2} - \frac{1}{2}$.

$$\frac{\partial \pi_{ij}}{\partial \theta_k} = \frac{1}{\sigma^2} \pi_{ij} \left(t_{ij} d_{jk} - \sum_{J=1}^{2^m} \pi_{iJ} t_{iJ} d_{Jk} \right)$$

$$\frac{\partial \pi_{ij}}{\partial \mu} = \frac{1}{\sigma^2} \pi_{ij} \left(t_{ij} - \sum_{J=1}^{2^m} \pi_{iJ} t_{iJ} \right)$$

$$\frac{\partial \pi_{ij}}{\partial \sigma^2} = \frac{1}{\sigma^2} \pi_{ij} \left(s_{ij} - \sum_{J=1}^{2^m} \pi_{iJ} s_{iJ} \right)$$

The following refer to $k = 1, \dots, c$ and $p = 1, \dots, c$.

$$\begin{aligned} \text{(1)} \quad \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_p \partial \theta_k} &= \frac{\partial}{\partial \theta_p} \left[\sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \pi_{ij} (y_i - \mu_j) d_{jk} \right] \\ &= - \sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \pi_{ij} d_{jk} d_{jp} + \sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \frac{\partial \pi_{ij}}{\partial \theta_p} t_{ij} d_{jk} \\ &= - \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} d_{jk} d_{jp} + \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} t_{ij}^2 d_{jk} d_{jp} \\ &\quad - \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \left[\left(\sum_{j=1}^{2^m} \pi_{ij} t_{ij} d_{jk} \right) \left(\sum_{J=1}^{2^m} \pi_{iJ} t_{iJ} d_{Jp} \right) \right] \end{aligned}$$

In matrix notation, we have:

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_p \partial \theta_k} &= - \frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} (\mathbf{D}^{[k]} \# \mathbf{D}^{[p]}) + \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) (\mathbf{D}^{[k]} \# \mathbf{D}^{[p]}) \\ &\quad - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[k]}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[p]}] \right\} \end{aligned}$$

where $\mathbf{1}_{1 \times n}$ is the $1 \times n$ matrix whose entries are all equal to 1.

(2) $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \theta_k}$ is composed of expressions similar to those in (1) except that the d_{jp} (and d_{Jp}) factors are replaced by 1.

In matrix notation, we have:

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \theta_k} &= - \frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{D}^{[k]} + \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) \mathbf{D}^{[k]} \\ &\quad - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[k]}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \right\} \end{aligned}$$

where $\mathbf{1}_{2^m \times 1}$ is the $2^m \times 1$ matrix whose entries are all equal to 1.

(3) $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \mu}$ is composed of expressions similar to those in (2) except that the d_{jk} (and d_{Jk}) factors are replaced by 1.

In matrix notation, we have:

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \mu} &= -\frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{1}_{2^m \times 1} + \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) \mathbf{1}_{2^m \times 1} \\ &\quad - \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} \{[(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}]\} \end{aligned}$$

Note that $\mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{1}_{2^m \times 1}$ equals n , the sample size.

$$\begin{aligned} (4) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_k \partial \sigma^2} &= \frac{\partial}{\partial \theta_k} \left[\sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \pi_{ij} \left[\frac{(y_i - \mu_j)^2}{2\sigma^2} - \frac{1}{2} \right] \right] \\ &= -\sum_{i=1}^n \sum_{j=1}^{2^m} \left(\frac{1}{\sigma^2}\right)^2 \pi_{ij} t_{ij} d_{jk} + \sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \frac{\partial \pi_{ij}}{\partial \theta_k} s_{ij} \\ &= -\left(\frac{1}{\sigma^2}\right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} t_{ij} d_{jk} + \left(\frac{1}{\sigma^2}\right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} s_{ij} t_{ij} d_{jk} \\ &\quad - \left(\frac{1}{\sigma^2}\right)^2 \sum_{i=1}^n \left[\left(\sum_{j=1}^{2^m} \pi_{ij} s_{ij} \right) \left(\sum_{J=1}^{2^m} \pi_{iJ} t_{iJ} d_{Jk} \right) \right] \end{aligned}$$

In matrix notation, we have:

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_k \partial \sigma^2} &= -\left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[k]} + \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{T}) \mathbf{D}^{[k]} \\ &\quad - \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} \{[(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[k]}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}]\} \end{aligned}$$

(5) $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma^2}$ consists of expressions similar to those in (4) except that the d_{jk} (and d_{Jk}) factors are replaced by 1.

In matrix notation, we have:

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma^2} &= -\left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1} + \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{T}) \mathbf{1}_{2^m \times 1} \\ &\quad - \left(\frac{1}{\sigma^2}\right)^2 \mathbf{1}_{1 \times n} \{[(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}]\} \end{aligned}$$

$$\begin{aligned}
(6) \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} \left[\sum_{i=1}^n \sum_{j=1}^{2^m} \frac{1}{\sigma^2} \pi_{ij} \left(\frac{(y_i - \mu_j)^2}{2\sigma^2} - \frac{1}{2} \right) \right] \\
&= - \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} (2s_{ij} + \frac{1}{2}) + \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{2^m} \frac{\partial \pi_{ij}}{\partial \sigma^2} s_{ij} \\
&= - \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} (2s_{ij} + \frac{1}{2}) + \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \sum_{j=1}^{2^m} \pi_{ij} s_{ij} s_{ij} \\
&\quad - \left(\frac{1}{\sigma^2} \right)^2 \sum_{i=1}^n \left[\left(\sum_{j=1}^{2^m} \pi_{ij} s_{ij} \right) \left(\sum_{J=1}^{2^m} \pi_{iJ} s_{iJ} \right) \right]
\end{aligned}$$

In matrix notation, we have:

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \sigma^2} &= - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# (2\mathbf{S} + \mathbf{H})) \mathbf{1}_{2^m \times 1} + \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{S}) \mathbf{1}_{2^m \times 1} \\
&\quad - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \{ [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \}
\end{aligned}$$

where \mathbf{H} is the $n \times 2^m$ matrix where every entry equals $\frac{1}{2}$.

The full log-likelihood second derivative matrix

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} &= \begin{pmatrix} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_1} & \cdots & \cdots & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_c} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \mu} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_1 \partial \sigma^2} \\ \vdots & \vdots & \vdots & \vdots & | & \vdots & | & \vdots \\ \vdots & \vdots & \vdots & \vdots & | & \vdots & | & \vdots \\ \hline \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_c \partial \theta_1} & \cdots & \cdots & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_c \partial \theta_c} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_c \partial \mu} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_c \partial \sigma^2} \\ \hline \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \theta_1} & \cdots & \cdots & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \theta_c} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \mu} & & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma^2} \\ \hline \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \theta_1} & \cdots & \cdots & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \theta_c} & | & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \mu} & & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix} \\
&= \begin{pmatrix} A_{c \times c} & X_{c \times 1} & Z_{c \times 1} \\ X'_{c \times 1} & & \\ & B_{2 \times 2} & \\ Z'_{c \times 1} & & \end{pmatrix}
\end{aligned}$$

where

$$A_{fg} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_f \partial \theta_g} = -\frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} (\mathbf{D}^{[f]} \# \mathbf{D}^{[g]})$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) (\mathbf{D}^{[f]} \# \mathbf{D}^{[g]}) - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[f]}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[g]}] \right\}$$

$$X_f = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \theta_g} = -\frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{D}^{[g]}$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) \mathbf{D}^{[g]} - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[g]}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \right\}$$

$$Z_g = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \theta_g \partial \sigma^2} = -\left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[g]}$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{T}) \mathbf{D}^{[g]} - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{D}^{[g]}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \right\}$$

$$B_{11} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \mu} = -\frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{1}_{2^m \times 1}$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T} \# \mathbf{T}) \mathbf{1}_{2^m \times 1} - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \right\}$$

$$B_{12} = B_{21} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma^2} = -\left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{T}) \mathbf{1}_{2^m \times 1} - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{T}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \right\}$$

$$B_{22} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2 \partial \sigma^2} = -\left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# (2\mathbf{S} + \mathbf{H})) \mathbf{1}_{2^m \times 1}$$

$$+ \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} (\boldsymbol{\Pi} \# \mathbf{S} \# \mathbf{S}) \mathbf{1}_{2^m \times 1} - \left(\frac{1}{\sigma^2} \right)^2 \mathbf{1}_{1 \times n} \left\{ [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \# [(\boldsymbol{\Pi} \# \mathbf{S}) \mathbf{1}_{2^m \times 1}] \right\}$$

Note that the first term in B_{11} , namely $-\frac{1}{\sigma^2} \mathbf{1}_{1 \times n} \boldsymbol{\Pi} \mathbf{1}_{2^m \times 1}$ equals $-\frac{1}{\sigma^2} n$.

Score Statistic

Assume that there are $c - 1$ parameters already in the model and that we are testing a new model with an additional parameter for a total of c parameters $\boldsymbol{\theta} = (\boldsymbol{\eta}, \beta) = (\theta_1, \dots, \theta_{c-1}, \beta, \mu, \sigma^2)$. Here β represents the 'new' c^{th} parameter and $\boldsymbol{\eta} = (\theta_1, \dots, \theta_{c-1}, \mu, \sigma^2)$ is considered the vector of nuisance parameters.

Let \widehat{U}_i denote the contribution of individual i to the approximated score function for the model where $\widehat{U} = \sum_{i=1}^n \widehat{U}_i$ is the approximated score function. (\widehat{U}_i is described below.) The score statistic for $H_0: \beta = 0$ against $H_1: \beta \neq 0$ is $W = \widehat{U}' \widehat{V}^{-1} \widehat{U}$ where $\widehat{V} = \sum_{i=1}^n \widehat{U}_i \widehat{U}_i'$.

We have that

$$\widehat{U}_i = U_{\beta,i}(0, \hat{\boldsymbol{\eta}}) - \left(\frac{\partial^2 l(0, \hat{\boldsymbol{\eta}})}{\partial \beta \partial \boldsymbol{\eta}} \right) \left(\frac{\partial^2 l(0, \hat{\boldsymbol{\eta}})}{\partial \boldsymbol{\eta}^2} \right)^{-1} U_{\boldsymbol{\eta},i}(0, \hat{\boldsymbol{\eta}})$$

where $\hat{\boldsymbol{\eta}}$ denotes the maximum likelihood estimator of $\boldsymbol{\eta}$ assuming c parameters in the model.

$$U_{\beta,i}(\beta, \boldsymbol{\eta}) = \frac{\partial l_i(\beta, \boldsymbol{\eta})}{\partial \beta} = \frac{1}{\sigma^2} (\boldsymbol{\Pi}_i \# \mathbf{T}_i) \mathbf{D}^{[c]}$$

$$\begin{aligned} U_{\boldsymbol{\eta},i}(\beta, \boldsymbol{\eta}) &= \frac{\partial l_i(\beta, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \left(\frac{\partial l_i}{\partial \theta_1}, \dots, \frac{\partial l_i}{\partial \theta_{c-1}}, \frac{\partial l_i}{\partial \mu}, \frac{\partial l_i}{\partial \sigma^2} \right)' \\ &= \frac{1}{\sigma^2} ((\boldsymbol{\Pi}_i \# \mathbf{T}_i) \mathbf{D}^{[1]}, \dots, (\boldsymbol{\Pi}_i \# \mathbf{T}_i) \mathbf{D}^{[c-1]}, (\boldsymbol{\Pi}_i \# \mathbf{T}_i) \mathbf{1}_{2^m \times 1}, (\boldsymbol{\Pi}_i \# \mathbf{S}_i) \mathbf{1}_{2^m \times 1})' \end{aligned}$$

Note that $U_{\beta,i}(\beta, \boldsymbol{\eta})$ is a scalar and $U_{\boldsymbol{\eta},i}(\beta, \boldsymbol{\eta})$ is an $(c + 1) \times 1$ matrix.

In the expressions below, A , Z , X , and B refer to the submatrices of the full log-likelihood second derivative matrix defined earlier.

$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \boldsymbol{\eta}}$ is a $1 \times (c + 1)$ matrix:

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \boldsymbol{\eta}} = (A_{1c}, \dots, A_{(c-1)c}, X_c, Z_c)$$

This assumes that β is the c^{th} parameter in the model. This can also be interpreted as the transpose of the c^{th} column of the full log-likelihood second derivative matrix

after deleting the entry in the c^{th} row.

$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\eta}^2}$ is a $(c+1) \times (c+1)$ matrix:

$$\begin{pmatrix} \tilde{A}_{(c-1) \times (c-1)} & \tilde{X}_{(c-1) \times 1} & \tilde{Z}_{(c-1) \times 1} \\ & \tilde{X}'_{(c-1) \times 1} & \\ & & B_{2 \times 2} \\ \tilde{Z}'_{(c-1) \times 1} & & \end{pmatrix}$$

\tilde{A} is the upper $(c-1) \times (c-1)$ submatrix of A , \tilde{X} consists of the first $c-1$ entries of W , and \tilde{Z} consists of the first $c-1$ entries of Z .

Brief Description of GEM-NR Algorithm

The GEM-NR method is an iterative process of updating parameter values in order to find maximum likelihood estimates for models with missing data. It is a hybrid of the generalized EM algorithm and the Newton-Raphson method (McLachlan & Krishnan 1997).

Suppose at step t , our parameter values are $\boldsymbol{\theta}^{(t)}$ and the current values of the mixing probabilities are denoted by $\pi_{ij}^{(t)}$. Note that $\pi_{ij}^{(t)}$ is computed from the current values $\mu_j^{(t)}$ and $\sigma^{2(t)}$.

Expectation Step: Compute the conditional expected complete-data log-likelihood with respect to the conditional distribution of the missing data given the observed data and the current estimated parameter values $\boldsymbol{\theta}^{(t)}$:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) = \sum_{i=1}^n \sum_{j=1}^{2^m} \log [\phi(y_i : \mu_j, \sigma^2) p_{ij}] \times \pi_{ij}^{(t)}$$

Maximization Step: Update $\boldsymbol{\theta}^{(t)}$ to $\boldsymbol{\theta}^{(t+1)}$ by:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \alpha^{(t)} \left[\frac{\partial^2 Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right]^{-1} \bigg|_{\boldsymbol{\theta}^{(t)}} \frac{\partial Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta}^{(t)}}$$

The appropriate partial derivatives can be found from the partial derivatives of the full log-likelihood above by setting the partial derivatives of π_{ij} equal to zero.

Choice of step-size $\alpha^{(t)}$

The parameter $\alpha^{(t)}$ can be chosen between 0 and 1. It is usually chosen to be 1. However, we need to ensure at each step that the likelihood increases. If the likelihood decreases, the step-size is decreased and the update recomputed. McLachlan & Krishnan (1997) describe a condition to check to ensure that the conditional log-likelihood and thus the likelihood increases. In our simulations, we directly check whether the likelihood increases at each iteration.

TABLE 1
QTL positions and effect sizes for three simulation cases

QTL	1	2	3	4	5	6	7	8
Case 1								
Chromosome	1	2	3	5	6	7	8	9
Positions (cM)	27.4	49.0	32.5	90.3	9.3	70.7	88.9	63.2
Additive effect	0.577	0.577	0.577	0.577	0.577	0.577	0.577	0.577
Case 2								
Chromosome	1	1	2	3	3	6	7	9
Positions (cM)	27.4	90.3	49.0	32.5	88.9	9.3	70.7	63.2
Additive effect	0.503	0.670	0.789	0.590	-0.503	0.710	0.255	0.399
Case 3								
Chromosome	1	1	1	2	2	3	3	3
Positions (cM)	7.6	52.3	99.0	32.5	88.9	9.3	41.7	83.2
Additive effect	0.503	0.670	0.789	0.590	-0.503	0.710	0.255	0.399

TABLE 2
FDR for Case 1, 2 and 3 with different significance levels
and LOD-support intervals

	α	5%	10%	15%	20%	25%	30%
Case 1	LOD-1.0	.090	.094	.100	.105	.110	.117
Case 2		.099	.107	.107	.119	.127	.134
Case 3		.132	.124	.123	.121	.123	.123
Case 1	LOD-1.5	.042	.046	.051	.056	.061	.069
Case 2		.045	.052	.057	.063	.071	.077
Case 3		.070	.064	.065	.066	.068	.069
Case 1	LOD-2.0	.023	.028	.033	.038	.042	.049
Case 2		.023	.029	.034	.038	.043	.047
Case 3		.036	.034	.035	.036	.040	.041

TABLE 3

Power to detect QTL for Case 2 for selected significance levels
(using LOD-1.5 support intervals)

QTL	Effect	α					
		5%	10%	15%	20%	25%	30%
1	.503	.630	.712	.766	.797	.809	.829
2	.670	.921	.943	.954	.960	.961	.967
3	.798	.973	.980	.979	.978	.977	.977
4	.590	.717	.800	.824	.844	.859	.868
5	-.503	.577	.668	.735	.758	.781	.795
6	.710	.968	.973	.976	.978	.979	.978
7	.255	.165	.221	.259	.295	.321	.353
8	.399	.461	.551	.613	.652	.684	.707

TABLE 4
Average position estimates, LOD coverage, and LOD width
for Case 2 at 20% significance level
(Standard deviations of position estimates are in parentheses)

QTL	Effect	Position (Chr:cM)	Estimate (SD)	LOD Coverage (%)			LOD Width (cM)		
				1	1.5	2	1	1.5	2
1	.503	1 : 27.4	27.3 (7.61)	88.8	95.4	97.8	20.9	29.2	38.7
2	.670	1 : 90.3	89.7 (4.34)	92.4	96.8	98.3	13.4	17.9	22.4
3	.798	2 : 49.0	49.1 (3.62)	92.8	95.4	96.9	11.9	15.4	19.1
4	.590	3 : 32.5	31.5 (6.18)	91.4	97.0	98.3	17.7	24.4	32.2
5	-.503	3 : 88.9	89.3 (6.53)	92.1	96.4	98.1	18.0	24.9	33.9
6	.710	6 : 9.3	9.9 (4.00)	92.3	96.1	97.2	12.7	16.2	20.1
7	.255	7 : 70.7	68.0 (17.28)	68.9	82.9	92.4	25.8	39.5	59.7
8	.399	9 : 63.2	62.8 (12.86)	81.8	90.4	95.8	24.4	37.4	54.0

TABLE 5

Average effect estimates, observed standard deviation over replicates, and
average estimated standard deviation
using observed Fisher information matrix
(Case 2, 20% significance level)

QTL	Chrom.	Effect	Eff. Estimate	Observed SD	Estimated SD
1	1	.503	.553	.122	.119
2	1	.670	.711	.141	.118
3	2	.798	.814	.140	.114
4	3	.590	.629	.139	.118
5	3	-.503	-.572	.125	.118
6	6	.710	.732	.131	.114
7	7	.255	.450	.094	.113
8	9	.399	.480	.088	.113

FIGURE LEGENDS:

Figure 1: Comparison of score statistic and associated threshold with likelihood ratio statistic and permutation threshold for Case 1. Figure 1A compares threshold values (y-axis) across significance levels α (x-axis) with score thresholds indicated by the dotted curve and permutation thresholds indicated by the solid curve. Figure 1B compares the likelihood ratio profile (solid) and score statistic profile (dotted) for one replication; the solid and dotted horizontal lines represent the permutation threshold and score threshold ($\alpha = 5\%$), respectively.

Figure 2: Score threshold as QTL are added to the model (average over replicates). The x-axis represents the number of QTL in the model being tested. Each line represents a significance level α : from top to bottom, $\alpha = 0.01, 0.02, 0.03, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30$.

Figure 3: Model size results: x-axis represents the number of QTL in the final model and the y-axis represents the percentage of replicates resulting in the given model size. For each case the simulated model contains 8 QTL. Figure 3A is for Case 1 (equal effects, no linkage, 9 chromosomes). Figure 3B is for Case 2 (varying effects, mild linkage, 9 chromosomes). Figure 3C is for Case 3 (varying effects, more linkage, 3 chromosomes). Percentages in the legend refer to significance levels α .

Figure 1A

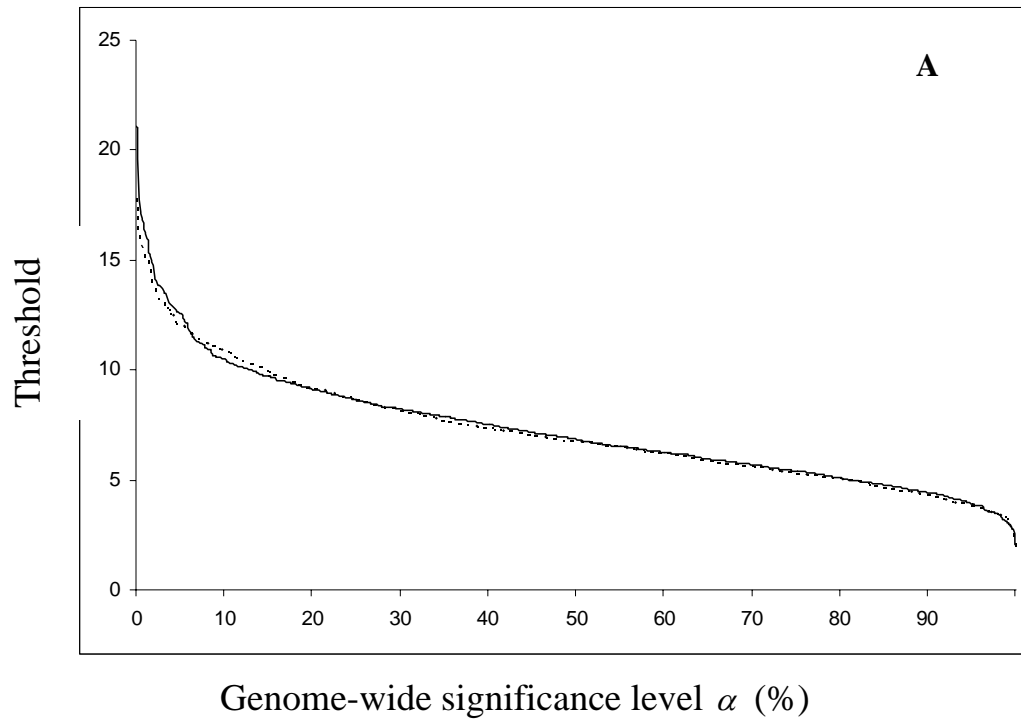


Figure 1B

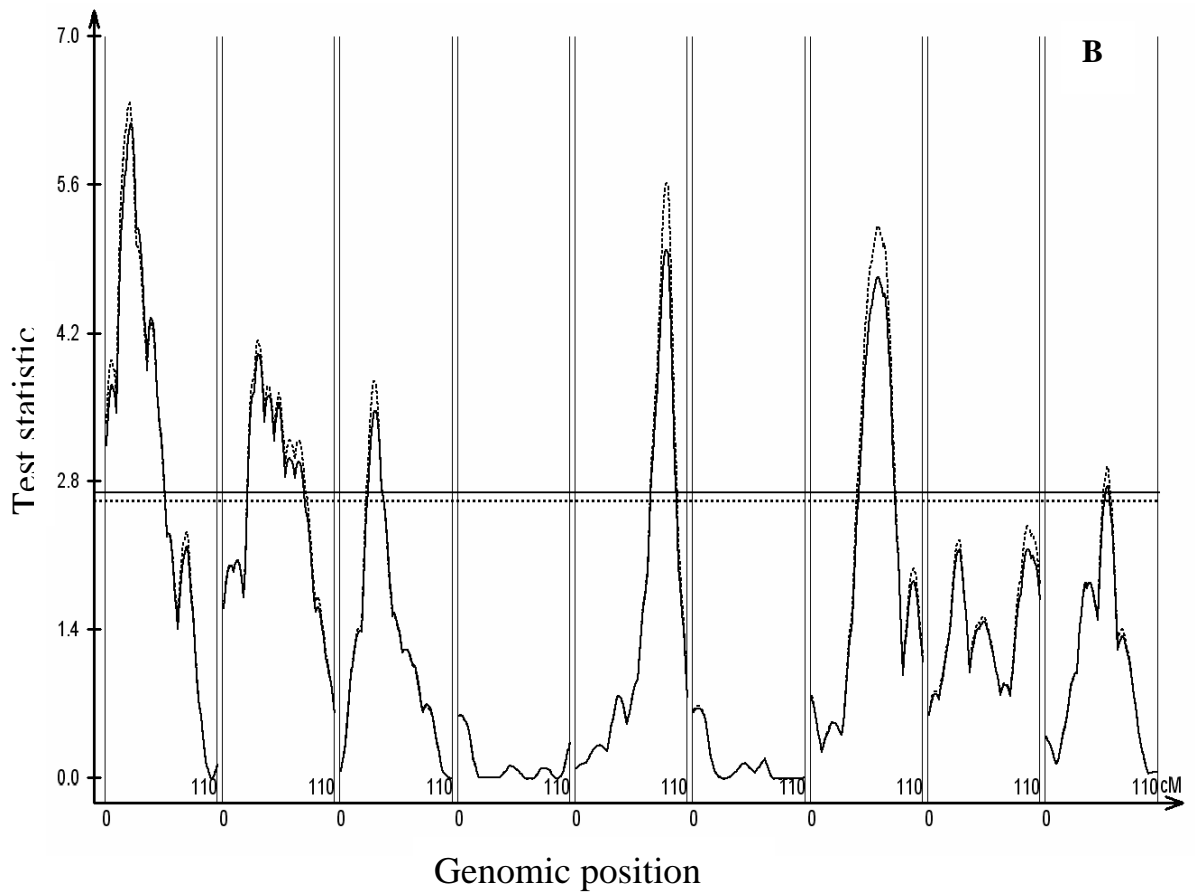


Figure 2

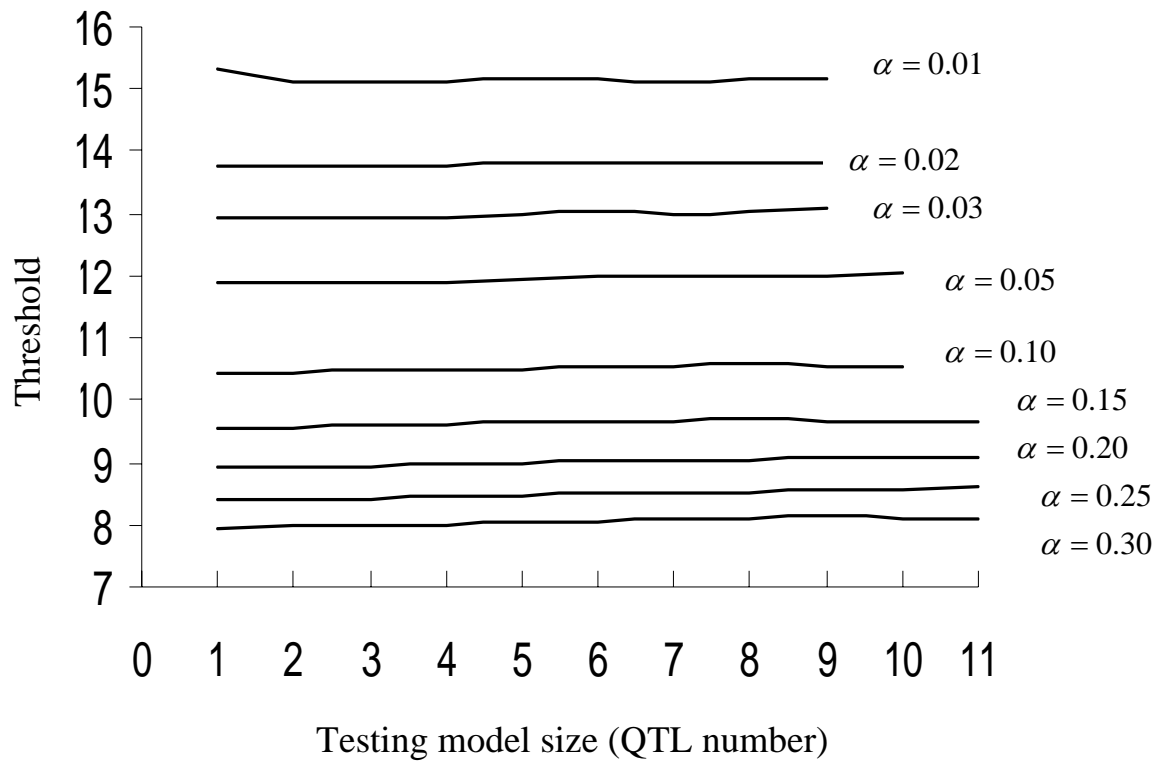


Figure 3

