Published in final edited form as:

J Stat Comput Simul. 2014 ; 84(3): 644–662. doi:10.1080/00949655.2012.721885.

Fence Methods for Backcross Experiments

Thuan Nguyen^{a,*}, Jie Peng^b, and Jiming Jiang^b

^aDepartment of Public Health and Preventive Medicine, Oregon Health and Science University, Portland, OR 97239, U.S.A.

^bDepartment of Statistics, University of California, Davis, CA 95616, U.S.A.

Abstract

Model search strategies play an important role in finding simultaneous susceptibility genes that are associated with a trait. More particularly, model selection via the information criteria, such as the BIC with modifications, have received considerable attention in quantitative trait loci (QTL) mapping. However, such modifications often depend upon several factors, such as sample size, prior distribution, and the type of experiment, e.g., backcross, intercross. These changes make it difficult to generalize the methods to all cases. The *fence method* avoids such limitations with a unified approach, and hence can be used more broadly. In this paper, this method is studied in the case of backcross experiments throughout a series of simulation studies. The results are compared with those of the modified BIC method as well as some of the most popular shrinkage methods for model selection.

Keywords

high-dimensional variable selection; restricted fence (RF); model selection

1. Introduction

Unraveling the genetic influence on phenotypic differences in human is often difficult due to the genetic and cultural heterogeneity of the populations. One approach to the problem has been to use appropriate animal models to pinpoint candidate genes for more focused further investigations. For example, the mouse model has been extensively used in quantitative trait loci (QTL) mapping. Mouse is relatively inexpensive to maintain and breed. In addition, it has a high rate of reproduction. The importance of mice in genetic studies has been established through the work of Lucien Cuénot, a French geneticist, who demonstrated Mendelian inheritance in mammals using the inheritance of coat colors in mice [1]. Castle & Allen often earned a credit for this success from their work published in 1903 [2]. Lucien Cuénot later on also demonstrated that a gene have multiple alleles [3] and that some alleles can be lethal, using the yellow allele of the agouti gene *A*^y[4]. Castle & Little confirmed this observation through their work and gave an explanation in 1910 [5]. Nowadays, developed mouse strains serve as primary models for many studies in human diseases, such as obesity, diabetes and cancer.

Ideally, the objective of the experiment is to identify susceptibility genes that contribute to the variation of a phenotype. However, in practice, such an ambitious aim is often reduced to identifying the genomic regions in which attributed genes are lying. Therefore, it is

desirable to form an association between the phenotypic variation and such informative genomic regions. Classical methods of QTL mapping includes interval mapping [6], composite interval mapping [7, 8], and multiple QTL mapping, or MQM [9, 10]. Recently, model selection methods via the information criteria, e.g., AIC [11], BIC [12] have received considerable attention in QTL mapping [13, 14, 16–21]. In particular, Broman & Speed [14] proposed a modification of the BIC criterion, called BIC_{δ} to overcome the overestimation of the number of QTLs when using the traditional BIC. However, the modification depends on the sample size. Namely, the authors obtained an appropriate value of δ via extensive Monte Carlo simulations, whose value changes with the sample size. In other words, unless the sample size of the study coincides with one that has already been considered, the Monte Carlo simulations have to be carried out again in order to determine the value of δ . Bogdan *et al.* [20] proposed another version of BIC modification, called mBIC, which includes an extra penalty term that depends on the number of markers and the choice of prior on the QTL numbers. Baierl *et al.* [21] modified the mBIC by adjusting its penalty and applied the method to intercross experiments. In short, all of the BIC modifications depend on several factors, such as sample size, prior distribution, and the type of experiment; thus, it is not easy to adopt a unified approach along these lines that can apply to all the cases.

The fence method [23, 24] was motivated by the limitations of the information criteria when applied to non-conventional situations, even though the method has been shown to be very competitive to traditional methods in the conventional settings as well. In particular, the adaptive fence procedure is a data driven procedure to determine an optimal tuning parameter. In terms of QTL mapping, once a parsimonious model is determined, the corresponding QTL number and locations can be identified. This method does not suffer from the limitations of the above mentioned methods based on BIC modifications. For example, although it is beyond the scope of this work, the method is potentially useful to human genetics, in which selection of the variance components can be essential.

It should be pointed out that there are a number of differences between the model-selection approach to QTL mapping and the traditional regression variable selection. Firstly, the regressors used in the QTL mapping, namely, the genotype indicators corresponding to the markers, are usually correlated (typically a Markov chain). Such correlations are expected to have an impact on the variable selection (here the variables correspond to the markers) in that if a marker is identified as a QTL, its neighboring markers also have a (good) chance to be selected, due to the correlation, especially when the distance between markers is small. Secondly, the signals at the true QTLs are typically fairly weak, for example, heritability of 50% or less. Finally, the number of potential markers to be examined is typically fairly large, leading to a large number of evaluations for the all-subset selection. For example, Broman & Speed [14] considered QTL mapping in a backcross experiment with 9 chromosomes, each with 11 markers. This led to a (conditional) linear regression model with 99 candidate variables, one corresponding to each marker. In fact, these difficulties have led the authors to modify the traditional BIC criterion for regression variable selection in order to apply it to QTL mapping problems (see below for more details).

We similarly treat the QTL mapping as a model selection problem, but our approach is based on the fence method, which is attractive in this situation due to its flexibility and data-driven optimality [23, 24]. On the other hand, the fence, especially the adaptive fence, encounters computational difficulties when applied to QTL mapping due to the potentially large number of markers, as mentioned above. For example, in the aforementioned backcross experiment considered in [14], the fence would require up to 2^{99} evaluations of the measures of lack-of-fit, and the adaptive fence would need to do the same for every bootstrap sample.

To overcome this computational difficulty, we propose to use a variation of the fence, known as the restricted fence (RF; [25]) for QTL mapping. We focus on the backcross experiment for the sake of simplicity, with a variety of settings that include different QTL numbers and locations, on markers or between flanking markers. In Section 2 we introduce a statistical model for the QTL mapping and show how to implement the RF to the current situation. A numerical algorithm is also provided. The performance of the RF is illustrated by a series of simulation studies in Section 3 where we also compare the RF with the BIC_{δ} method. We conclude with a discussion in Section 4 that includes additional simulation results regarding comparisons with the shrinkage methods for variable selection.

2. A statistical model, method, and algorithm

2.1. A Statistical model for backcross experiments

In human and mouse during meiosis, crossovers may occur on autosomal chromosomes. This recombination process mixes up the genetic materials that are passed from the parents to their offspring. As a result, the genetic contribution that the offspring have inherited from their ancestors is a random mosaic of genetic segments. In experimental studies, the segregation of a trait can be studied using the crosses between inbred strains, e.g., backcross or intercross. These are the standard approaches for dissecting an heritable trait that is based upon the cross of two inbred strains which show a large difference in their phenotypic values. More specifically, an outcross between these two inbred strains will be performed to obtain the first filial generation, or F_1 . All individuals within the same inbred strain are genetically identical and homozygous at all loci throughout the genome, hence the F_1 offsprings are heterozygous at all loci at which the parental strains differ. The choice of which of the two original inbred strains to use depends on the variability of the phenotype. Larger phenotypic variability usually means larger heritability, h^2 , hence a larger chance of detecting QTLs. The backcross is obtained by mating the F_1 offspring with either one of the original inbred strains. As a result, there are two possibilities, QQ (homozygous) and Qq (heterozygous) genotypes, in a backcross design [13, 26].

Let y_i denote the phenotypic value of individual i , and let $x_{ij} = 1$ (homozygote) or $x_{ij} = 0$ (heterozygote) be the genotype of individual i at marker j . It is assumed that, within the same chromosome, the x_{ij} 's follows a Markov chain model with transition probabilities

$$P(x_{i,j+1}=1|x_{ij}=0) = P(x_{i,j+1}=0|x_{ij}=1) = \theta_j,$$

where θ_j is the *recombination fraction* between markers j and $j + 1$; and that $P(x_{ij} = 1) = P(x_{ij} = 0) = 1/2$ in accordance with Mendel's rules. As pointed out by Broman & Speed ([14]), the QTL identification may be viewed as a model selection problem. Considering the case where QTLs are on the markers, these authors employed an additive model, conditional on the genotypes:

$$y_i = \mu + \sum_{j \in S} \beta_j x_{ij} + \epsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where S is a subset of the indexes corresponding to the marker regressors. The errors, ϵ_i , are assumed to be independent and distributed as $N(0, \sigma^2)$, where σ^2 is an unknown variance.

We also extend our investigation to the case in which QTLs are no longer at markers. In other words, only markers near but not exactly at functional polymorphisms are genotyped; more specifically, the QTLs are located in the middle of their flanking markers. Consider a simple linear model to test for a QTL located between markers j and $j + 1$

$$y_i = b_0 + b^* x_i^* + \epsilon_i, \quad i = 1, 2, \dots, n,$$

where b^* is the effect of the putative QTL expressed as a difference in effects between the homozygote and heterozygote, x_i^* is an unobservable indicator variable, taking a value 1 or 0 depending on the genotype of the QTLs under consideration. Given the genotypes of the two flanking markers j and $j + 1$, one can calculate the probability of x_i^* being one or zero. Namely (e.g., Lynch & Walsh 1998, ch.15, [27])

$$\begin{aligned} P(x_i^* = 1 | x_j = 1, x_{j+1} = 1) &= P(x_i^* = 0 | x_j = 0, x_{j+1} = 0) = \frac{(1-\theta_1)(1-\theta_2)}{1-\theta}; \\ P(x_i^* = 1 | x_j = 0, x_{j+1} = 1) &= P(x_i^* = 0 | x_j = 1, x_{j+1} = 0) = \frac{\theta_1(1-\theta_2)}{\theta}; \\ P(x_i^* = 0 | x_j = 0, x_{j+1} = 1) &= P(x_i^* = 1 | x_j = 1, x_{j+1} = 0) = \frac{(1-\theta_1)\theta_2}{\theta}; \\ P(x_i^* = 1 | x_j = 0, x_{j+1} = 0) &= P(x_i^* = 0 | x_j = 1, x_{j+1} = 1) = \frac{\theta_1\theta_2}{1-\theta}, \end{aligned}$$

where, θ_1, θ_2 are the recombination fractions between markers $j, j + 1$ and the putative QTL, and $1-\theta = \theta_1\theta_2 + (1-\theta_1)(1-\theta_2)$. In particular, when recombination fraction θ between markers j and $j + 1$ is small, we have $\theta \approx \theta_1 + \theta_2$. More specifically, $\theta_1 = \theta_2 = \theta/2$ when the QTL is in the middle of the two flanking markers. The performances of the restricted fence and BIC _{δ} methods are, again, examined through simulation studies. Furthermore, because only the markers are considered as regressors in the class of candidate models, the true model, that is, model that includes all the true QTLs and nothing else, is not among the ones in model space. Therefore, we use the closest approximation as the basis for our evaluations of these methods. Some further results are deferred to Appendix.

2.2. Restricted fence method

A key issue in statistical model selection is how to strike a balance between model fit and complexity. This is an issue because, with a sufficiently complex model, one can expect that the model will fit well to the data. However, this does not mean that such a model is very useful, because other aspects of the model, such as simplicity, also need to be considered. The traditional information criteria, e.g., AIC [11], BIC [12], handle the balance between model fit and complexity by a criterion function of the form $c(M) = \hat{Q}(M) + \lambda_n |M|$, where M represents a candidate model, $\hat{Q}(M)$ is a measure of lack-of-fit in the sense that the value of $\hat{Q}(M)$ decreases as the complexity of M increases [the detailed definition of $\hat{Q}(M)$ is given below]. Furthermore, $|M|$ denotes the dimension of M in terms of the number of free parameters; thus $|M|$ increases as the complexity of M increases. Finally, λ_n is a penalty to the model complexity, which may depend on the sample size n . The optimal model is the one that minimizes $c(M)$ among all candidate models M . It is seen that the information criteria attempt to balance the model fit, and model complexity, by considering a weighted sum of $\hat{Q}(M)$ and $|M|$, where the weight is determined by the penalty λ_n .

Despite the popularity of the information criteria, it is known that these criteria may not work well when applied to non-conventional situations. For example, Broman & Speed [14] found that the BIC, for which $\lambda_n = \log(n)$, is under-penalizing when applied to backcross experiments. As a consequence, the method tends to include many extraneous markers, or false discoveries. The authors proposed a modified version of BIC, called BIC _{δ} and showed that it works much better (see below for more detail). Jiang *et al.* [23] introduced another way of balancing the model fit and complexity, known as *fence methods*. The authors noted that the issue is similar, in a way, to the type I and type II errors in hypothesis testing, in which one type of errors goes up as the other type goes down. It is customary, in hypothesis

testing, to control the type I error by the level of significance, α . One then tries to reduce the type II error given that the type I error is bounded by α . The fence uses a similar idea to control the model fit, in terms of an upper bound for $\hat{Q}(M)$. Given that $\hat{Q}(M)$ is within the upper bound, the fence tries to find the least complex model. The upper bound for $\hat{Q}(M)$ is called a fence, or barrier. A model is within the fence if its $\hat{Q}(M)$ is less than or equal to the upper bound. Once the fence is constructed, the optimal model is selected from those within the fence according to a criterion which can incorporate quantities of practical interest. Here, we consider simplicity of the model in terms of minimum $|M|$ as the only criterion to select a model within the fence. Using the mathematical notation, the fence is constructed via the following inequality

$$\hat{Q}_M - \hat{Q}_{\tilde{M}} \leq c_n \hat{\sigma}_{M, \tilde{M}}, \quad (2)$$

where $Q_M = Q_M(y, \theta_M)$ is a measure of lack-of-fit, y represents the vector of observations, M indicates a candidate model, and θ_M denotes the vector of parameters under M . Here by lack-of-fit we mean that Q_M satisfies the basic requirement that $E(Q_M)$ is minimized when M is a correct model, and θ_M the true parameter vector under M . Furthermore,

$\hat{Q}_M = \inf_{\theta_M \in \Theta_M}$ (the minimum value of Q_M over all θ_M that belong to Θ_M), where Θ_M is the parameter space under M , and $\tilde{M}$ is a model that minimizes $\hat{Q}_M$ among $M \in \mathcal{M}$, the set of candidate models. Finally, $\hat{\sigma}_{M, \tilde{M}}$ is an estimate of the standard deviation of $\hat{Q}_M - \hat{Q}_{\tilde{M}}$. The tuning constant c_n on the right side of (2) can be chosen using an adaptive procedure.

The calculation of $\hat{Q}_M$ is often straightforward. For example, in many cases Q_M can be chosen as the negative log-likelihood or residual sum of squares (RSS). On the other hand, the computation of $\hat{\sigma}_{M, \tilde{M}}$ can be quite challenging. Even if an expression can be obtained for $\hat{\sigma}_{M, \tilde{M}}$, its accuracy as an estimate of the standard deviation cannot be guaranteed in a finite sample situation. For such a reason, this step of the fence method has complicated its applicability to many areas. Jiang *et al.* [24] developed a simplified adaptive fence (SAF) procedure that avoids such difficulties. In the SAF procedure, the fence inequality (2) is replaced by

$$\hat{Q}_M - \hat{Q}_{\tilde{M}} \leq c_n, \quad (3)$$

It appears that the only difference is the disappearance of $\hat{\sigma}_{M, \tilde{M}}$ from the right side of (2). In fact, this term is merged into the tuning constant c_n , which is then chosen adaptively in the same way as in [23]. In [24], the SAF is shown to be consistent under suitable regularity conditions, and have outstanding finite sample performance.

However, even with the SAF procedure, one may still encounter computational difficulties when applying the fence method to high dimensional problems as is often the case in QTL mapping. To overcome this computational difficulty, Nguyen & Jiang [25] proposed the following variation of the SAF, known as the restricted fence (RF). The idea may be viewed as a combination of the idea of the restricted maximum likelihood (REML; e.g., [22]) and the SAF. Firstly, we apply a transformation to the data that is orthogonal to a (large) subset of candidate variables to make them not participate in the model. The SAF is then applied to the remaining (small) subset of variables. The term “restricted” is used because the first step of the proposed procedure involves the same transformation of the data as in REML (e.g., [22], p. 13); however, there is no estimation of the variance components. We show how to implement the RF method to the backcross experiment using an example.

Broman & Speed [14] considered a model selection problem in QTL mapping in a backcross experiment, which is a special case of section 2.1 with 9 chromosomes, each with 11 markers. This led to a (conditional) linear regression model (1) with 99 candidate variables, one corresponding to each marker. Write $X_1 = (x_{ij})_{1 \leq i \leq n, j \in S_1}$, and $X_2 = (x_{ij})_{1 \leq i \leq n, j \in S_2}$, where S_1 is a subset of S , and $S_2 = S \setminus S_1$. For example, S_1 may correspond to the subset of markers belonging to the first half of the first chromosome, and S_2 the rest of the markers. Then the model can be expressed, in matrix form, as $y = X\beta + \epsilon = X_1\beta^{(1)} + X_2\beta^{(2)} + \epsilon$, where $y = (y_i)_{1 \leq i \leq n}$, $\beta = (\beta_j)_{j \in S}$, $\beta^{(1)} = (\beta_j)_{j \in S_1}$, $\beta^{(2)} = (\beta_j)_{j \in S_2}$, and $\epsilon = (\epsilon_i)_{1 \leq i \leq n} \sim N(0, \sigma^2 I_n)$. Let $p_j = \text{rank}(X_j)$, $j = 1, 2$. Let A be a $n \times (n-p_2)$ matrix such that $A'A = I_{n-p_2}$, $A'X_2 = 0$. It follows that $AA' = P_{X_1} = I_n - P_{X_2}$, where $P_{X_2} = X_2(X_2'X_2)^{-1}X_2'$. Then, we have $z = A'y = \tilde{X}_1\beta_1 + \eta$, where $\tilde{X}_1 = A'X_1$, and $\eta = A'\epsilon \sim N(0, \sigma^2 I_{n-p_2})$.

Note that, by applying the transformation A' to the data, the matrix X_2 , which is often chosen to be of much higher dimension than X_1 , is opted out of the model. Thus, one can apply the SAF to the subset of markers corresponding to X_1 , which is usually in much lower dimension, using $z = A'y$ as the data. Also note that, the explicit form of the transformation matrix A is usually not needed for the application of the fence method. For example, if Q_M is chosen as RSS, the residual sum of squares, then it can be shown that the measure of lack-of-fit based on $z = A'y$ is

$$\hat{Q}_M = y' P_{X_1 \oplus X_2} y \quad (4)$$

with $P_{X_1 \oplus X_2} = P_{X_1} - P_{X_1} X_1' (X_1' P_{X_1} X_1)^{-1} X_1' P_{X_1}$ (see [25]), and the estimator of β_1 is given by

$$\hat{\beta}_1 = (\tilde{X}_1' \tilde{X}_1)^{-1} \tilde{X}_1' z = (X_1' P_{X_1} X_1)^{-1} X_1' P_{X_1} y. \quad (5)$$

Furthermore, for the SAF procedure, one can bootstrap under the full model restricted to S_1 without having to know or estimate β_2 . In fact, the bootstrap version of $\hat{Q}_M$ is given by

$$\hat{Q}_M^* = (X_1 \hat{\beta}_1 + \epsilon^*)' P_{X_1 \oplus X_2} (X_1 \hat{\beta}_1 + \epsilon^*), \quad (6)$$

where ϵ^* is the vector of bootstrapped errors [25].

In Section 3, we apply the RF to each half chromosome. The SAF is then applied to all the markers that are picked up by the RF (from all of the half-chromosomes) in order to identify the final QTLs. The detailed steps are given below as a numerical algorithm for the restricted fence procedure.

2.3. Algorithm

1. For the candidate variables x_j , $1 \leq j \leq J$, determine a division $S = \{1, \dots, J\} = \mathcal{T}_1 \cup \dots \cup \mathcal{T}_q$, where $\mathcal{T}_r$, $1 \leq r \leq q$ are subsets of S (not necessarily disjoint).
2. Let $S_1 = \mathcal{T}_1$ and $S_2 = S \setminus S_1$. Apply the SAF (stage 1) using the measure of lack-of-fit (4) to select the variables among x_j , $j \in S_1$. The SAF consists of the following steps:

- 2.1. Estimating the parameters under the restricted full model ($= \{x_j, j \in S_1\}$).
- 2.2. Bootstrapping under the restricted full model. Let $c^{(1)} \dots c^{(K)}$ be a grid of values of c_n being considered such that $0 < c^{(1)} < \dots < c^{(K)}$. For each bootstrapped sample, select the best model using the fence (3), for each of the c_n values among the grid.
- 2.3. For each of the c_n values among the grid, compute the frequency, over the bootstrap samples, that each candidate model is selected as the best model; compute the maximum frequency, denoted by p^* . Note that p^* depends on the value of c_n .
- 2.4. Find a peak in the middle of the plot of p^* against c_n (over the grid); let c_n^* be the c_n corresponding to the peak; use the fence (3) with $c_n = c_n^*$ to select the final optimal model for the subset $\mathcal{T}_1$.
3. Apply the same procedure as 2 to $\mathcal{T}_2, \dots, \mathcal{T}_q$.
4. Apply another SAF (stage 2) to the subset of variables selected in 2 and 3 (combined; considered as the new candidate variables) to select the final variables.

Example. In the backcross experiment considered in [14], J is the total number of markers, that is, $J = 99$. We may let $\mathcal{T}_1$ be the subset of the first 6 markers on chromosome 1, $\mathcal{T}_2$ be the subset of the last 5 markers on chromosome 1, $\mathcal{T}_3$ be the subset of the first 6 markers on chromosome 2, and so on. Thus, $q = 18$ in this case. In step 2, $S_1 = \mathcal{T}_1$ consists of 6 markers, and S_2 is the rest of the $99 - 6 = 93$ markers. Then, in Step 3, $S_1 = \mathcal{T}_2$ and S_2 is the rest of the markers; thus, S_1, S_2 consist of 5 markers and $99 - 5 = 94$ markers, respectively, and so on.

3. Simulation studies

We carry out a number of simulation studies for the same backcross experiment considered in [14]. The number of progeny is chosen as $n = 500, 750$ or 1000 . Nine chromosomes are considered. It is assumed that the length of each chromosome is 100 cM. 11 equally spaced markers (at a spacing of 10 cM) are genotyped. The recombination process is assumed to be without crossover interference. The Haldane recombination function is applied to compute the recombination rate. The marker data are assumed to be complete and without errors. The number of underlying QTLs is either 7 or 9. In the 7-QTL case, equal effects, $|\beta| = .76$ at each QTL, are examined. The setting is similar to that of [14]. The first two QTLs are located at 4th and 8th markers of the 1st chromosome. These two QTLs have *coupling* link (i.e., their effects have the same sign). The next two QTLs are located at 4th and 8th markers of the 2nd chromosome, but have *repulsion* link (i.e., their effects have the opposite sign). The next three QTLs are located at 6th, 4th, and 1st markers of the 3rd, 4th, and 5th chromosomes, respectively. The last four chromosomes contain no QTL. The errors are normally distributed with mean 0 and $\sigma = 1$. As a result, the *heritability* of the trait which is the proportion of the variance attributed by genetic effects with respect to the total phenotypic variance is 50% (see the Appendix). In the 9-QTL case, the first 7 QTLs are the same as in the previous case, and the last two QTLs are located at the 1st and 6th markers of the 6th chromosome. These two new QTLs on the 6th chromosome have a *coupling* link. To maintain the *heritability* at 50%, the effects of these nine QTLs are reduced to .636. There are 100 simulation replicates in each case.

We compare the RF method with the BIC_δ method, proposed by Broman & Speed [14]. The latter authors modified the traditional BIC procedure by adding an additional tuning parameter δ to the penalty part. The method aims to minimize the criterion function:

$$\text{BIC}_{\delta}(M) = \log \{ \text{RSS}(M) \} + \delta |M| \frac{\log(n)}{n}, \quad (7)$$

where M denotes a candidate model, $\text{RSS}(M)$ is the residual sum of squares under model M , and $|M|$ is the dimension of M . Note that if $\delta = 1$, (7) becomes the BIC. Regarding the choice of δ , the authors suggested $\delta = 2L / \log_{10}(n)$, where L denotes a threshold, namely, the 95th percentile of the maximum of the LOD scores, across the whole genome, under the null hypothesis that there is no QTL, which is determined through Monte-Carlo simulations.

Due to the high dimensionality in QTL mapping, Broman & Speed [14] incorporated forward selection and backward elimination procedures with BIC_{δ} . Their most recommended procedure is forward/backward BIC_{δ} . They also suggested using 25% of the candidate variables as the stopping rule for the forward selection, and then performing backward elimination. This means, for example, in the case of 99 (markers) covariates, the forward selection will include variables sequentially until 25 markers are picked up, and then backward elimination is applied. The optimal model is selected as the one with the minimum BIC_{δ} . However, to avoid missing relevant markers since here the signals are expected to be weak, we instead use a 50% stopping rule in the simulation experiments.

3.1. QTL on the marker

Here we consider the case when putative markers are completely linked to the QTLs. As mentioned earlier, for complex traits, several susceptibility genes may influence phenotypic variation simultaneously, such that the effect of each gene is rather small. As a consequence, it is difficult to detect the QTLs. Implicitly, this problem emerges while applying the restricted fence procedure in finding the QTLs, especially in the case of moderate sample size. Therefore, we propose some practical adjustments while applying the fence methods to genetic applications in general.

Firstly, we smooth the plot of p^* vs. c_n to remove the noise due to the bootstrap sampling, where p^* is the highest empirical probability that a model is selected (see [24]). Here the smoothing function `loess()` in R is used. Figures 1 - 3 help to explain why such a smoothing step is helpful. For example, consider the row2-column2 plot of figure 1, without the smoothing, the small peak due to noise (near $c_n = 5$) would have been mistakenly identified as the first peak (see below). Figure 1 includes the plots of p^* vs. c_n of SAF (stage 1) in the steps 2 and 3 described in the algorithm. These nine plots are corresponding to the nine first halves of the nine chromosomes under the consideration. For example, the row1-column1 plot is the SAF applied to the first six markers of chromosome 1. As expected, we picked up one QTL (one peak in the plot) from this consideration. In the row1-column2 plot, the first six markers of the chromosome 2 are considered. Similarly, we expect one peak in the plot to confirm there is one QTL within this region; the 4th marker has a signal $\beta = .76$ in the simulation setting. The row1-column3 plot is the plot for the first five markers of the chromosome 3. Since there is no QTL in this region—none of these five markers has any signal (the β s are zero)—one expects no peak in this plot, which is the case here. Similar explanation could be applied for the other six plots in this Figure 1. Likewise, in Figure 2, the nine plots are for the nine second halves of the nine chromosomes being considered. The explanations of these plots are similar to those in Figure 1. Also in Figure 2, the plots row2-column3, and row3-column1 show a small peak that smoothing seems not to be able to remove. The reason is that the smoothing parameter has to be chosen a priori for the entire 100 simulation runs; thus, such encounters cannot always be avoided. Our goal is to use the smoothing to screen tiny peaks as many as possible. Still, with such small peaks, the SAF (stage 1) may pick up some extraneous markers, or false discoveries. Therefore, we need the

use of SAF one more time (stage 2) on the markers selected from the stage 1. The stage 2 SAF will help eliminate such false discoveries.

Secondly, instead of choosing c_n at which p^* is the highest, we select c_n based on the “first significant peak” criterion. This adjustment is necessary in the case of weak signals. As discussed in [23], with an extremely large value of c_n , the smallest model is always selected. Now move along the curve from the right to the left, equivalently from large to smaller c_n , every time some QTLs enter the model we would encounter a peak in the curve. Thus, it is reasonable to think of the leftmost peak, or first peak, as the place where the last QTL(s) is included into the model. This interpretation is supported by Figure 3. More specifically, Figure 3 includes some p^* vs. c_n plots of SAF (stage 2) in step 4 described in the algorithm. The upper left plot is the one without smoothing. The upper right plot is the same one with smoothing. As mentioned, smoothing helps remove false discoveries that could happen in some cases due to the noise from bootstrapping technique. In the lower left plot, the first significant peak is identified. Without using the “first significant peak” adjustment, this peak is mistakenly missed. The lower right plot shows the case where the highest peak is also the first peak. In fact, this is what one expects to see in the case of strong signal, or larger sample size. Unfortunately, it is rarely seen in the case of weak signal or moderate/small sample size. The plots for the 9-QTL case are very similar (except that there are more peaks), and therefore omitted.

Results obtained from the restricted fence method are compared with those of the BIC_δ method proposed by Broman & Speed [14]. The latter authors suggested the values of δ as 2.56, 2.1, and 1.85 for $n = 100, 250$, and 500 , respectively. We obtained the similar values (using the same number of simulation runs - 50,000) at the 97.5th percentile threshold. To our knowledge, the LOD scores rely on the asymptotic distribution of the likelihood ratio statistic which has an asymptotic χ^2 distribution (not the t-distribution). Therefore, the error rate, i.e., $\alpha = 0.05$, should be used for the one-tail alternative. In other words, the 95th percentile threshold would be more accurate than the 97.5th percentile threshold. In fact, we obtained δ values at the 95th percentile as 2.23, 1.85, 1.64 for $n = 100, 250$, and 500 , respectively. The results based on these δ values have shown better performance in our simulation studies (see tables 1-4). We also obtained the value of δ for the cases of $n = 750$ and $n = 1000$ for our simulation settings. Since the δ values in these two cases are very close to each other, after rounding off (for $n = 750$) and rounding up (for $n = 1000$) with two decimal numbers, they end up with the same values, 1.7 and 1.48 for 97.5th and 95th percentiles, respectively.

We evaluate the performance of each procedure based on three criteria: (i) %TP denotes the empirical probability, in percentage, of identifying all true QTLs and nothing else (it should be noted that the cases in which the procedure detects all true QTLs plus some extraneous markers, or false positives, do not count for % TP); (ii) M(SD)TP denotes the mean (standard deviation) of the number of the true QTLs that are detected; and (iii) M(SD)FP denotes the mean (standard deviation) of the number of false positives. The simulation results are summarized in Table 1. In the case of sample size $n = 500$, the RF is very competitive to the forward/backward BIC_δ for example, 82% vs. 84% by %TP. The average number of true positives is slightly better in BIC_δ compared to that of the RF, 6.82 vs. 6.89 in (ii) (the target is 7 QTLs). However, the average number of false positives in the RF is slightly less than that of the forward/backward BIC_δ , 0.16 vs. 0.19. As for $n=750$, the RF is almost perfect in all three criterion while the forward/backward BIC_δ is not reaching the same results comparatively; more specifically, 98% vs. 94% in (i); 6.98 vs. 7 in (ii); and 0.02 vs. 0.07 in (iii). As noticed, the gap of false discoveries is more than 3-fold larger for the forward/backward BIC_δ compared to the RF's. It appears that overfitting occurs in the forward/backward BIC_δ procedure. This results in a slightly better performance in (ii), yet

much worse in (iii). Following Broman & Speed's [14] extended criterion of QTL detection, that is, a marker is detected as QTL if it is within 10 cM (B.1 column, Table 1) or 20 cM (B.2 column, Table 1) from the true QTL, we also found that RF overall performs better than the forward/backward BIC_{δ} (see Table 1 for more details). Table 2 reports the results for the 9-QTL case and the observations are similar.

3.2. QTL in the middle of two flanking markers

So far in the simulation study an ideal situation has been considered, that is, the QTLs are genotyped. However, such a situation is hardly practical. Under a more realistic scenario, one would genotype markers that are only partially linked to the QTLs. In other words, only genetic markers near functional polymorphisms are genotyped. Note that the markers do not by themselves associate with the phenotypic variation. However, since the markers are linked to the functional polymorphism, which in turn is correlated with the phenotype, one may observe associations between the markers and the phenotype. Clearly, the chance of establishing an association between the markers and the trait depends on the strength of the linkage between the functional polymorphism and the markers. Thus, the most challenging situation would be when the QTLs reside in the middle of two markers. This is our focus in the current subsection.

The simulation settings are similar to those of the previous subsection except that the locations of QTLs are no longer at the markers. More specifically, each QTL lies on the midpoint of its two flanking markers. So, in the 7-QTL case, the first QTL is at the midpoint of the 4th and 5th markers of the 1st chromosome. Similarly, for the second QTL, its flanking markers are the 8th and 9th markers of the 1st chromosome, and so on. The 9-QTL case is similar. With the same effects (e.g., coefficients) as the previous subsection, the empirical *heritability* is 40% in both the 7-QTL and 9-QTL settings.

As mentioned above, the true model that includes all the true QTLs (lying in the middle of two flanking markers) is not among the model candidates being considered (the model candidates involve markers only). Thus, we evaluate the performance of our method (RF) and the comparing method, BIC_{δ} based on the closest approximating selected model. That is, a (putative) QTL is detected if a selected marker (from RF or BIC_{δ} procedure) is either one of the flanking markers of the particular QTL. The results are summarized in Table 3 for the 7-QTLs case. More specifically, in the case of $n = 500$, the BIC_{δ} performs slightly better than the RF in the first two criteria: 53% vs. 59% in (i); and 6.4 vs. 6.63 in (ii). The RF and BIC_{δ} perform very closely in (iii), 0.56 vs. 0.57, respectively. When the sample size increases, $n = 750$, The RF outperforms 76% vs. 53% in (i); 6.75 vs. 6.91 in (ii); and .27 vs. 0.6 in (iii), compared to BIC_{δ} . Again, RF performs much better and BIC_{δ} in (i) and (iii) in the larger sample size case. Similarly, the overfitting of BIC_{δ} leads to a slightly better result in (ii). As in the previous subsection, with the extended criterion of QTL detection (see columns B.1 and B.2 in Table 3), we found the false positives of the BIC_{δ} can be somewhere from 6 times to more than 20 times higher than those of the RF, e.g., 0.05 vs. 0.33 in column B.1, and 0.03 vs. 0.79 in column B.2.

Similarly, in Table 4 for the 9-QTLs case, we observe the following comparative results of RF vs. BIC_{δ} with $n = 750$: 58% vs. 60% in (i); 8.47 vs. 8.73 in (ii) (the target is 9 QTLs); and 0.47 vs. 0.53 in (iii). When the sample size increases to $n = 1000$, the corresponding results are: 75% vs. 59% in (i); 8.7 vs. 8.9 in (ii); and 0.3 vs. 0.53 in (iii). The comparative performance of the RF is even better with the extended criterion of QTL detection (see columns B.1 and B.2 for more details).

An important observation is that, unlike the previous case where the QTLs are on the markers, here the performance of BIC_{δ} does not seem to improve (actually, in most cases, it

gets worse) when n increases. On the other hand, the RF always performs better when the sample size increases.

4. Discussion and further simulation results

In this paper, we evaluate the performance of a model selection procedure based upon three criteria: (i) % of true model detection; (ii) mean number of true QTLs identified, or positive discoveries; (iii) mean number of false QTLs selected, or false discoveries. In the case of moderate sample size, i.e., $n = 500$ for the 7-QTL case, or $n = 750$ for the 9-QTL case, the RF performs competitively with the BIC_δ in the first and third criterion. However, BIC_δ performs slightly better than the RF in the second criterion. This is partially due to the overfitting of BIC_δ while the fence tends to be more conservative in picking up a putative QTL, especially in the situation of weak signal or moderate/small sample size. As a consequence, the RF does a much better job in the third criterion, the false discoveries. As the sample size increases, the RF performs much better in the first and third criteria, and very close in the second. These observations hold in both settings—QTLs on the markers, and QTL in the middle of two flanking markers.

We would like to remark that in the case where QTLs are in the middle of flanking markers, the performance of the RF method improves as the sample size increases, while that of the BIC_δ method does not; in fact, when n increases from 500 to 750 in the 7-QTL case, or 750 to 1000 in the 9-QTL case, the performance of BIC_δ gets worse (see Tables 3 & 4).

The current work focuses on additive models in the backcross experiment. This means that only main effects are examined. On the other hand, epistatic effects are known to play important roles in many traits [17, 28–31]. This would result in a very high dimensional problem if all the interactions are taken into account simultaneously. For example, in our simulation setting where there are 99 markers (main effects), if all the two-way interaction terms are included, the dimension would go up to 4950, hence $p \gg n$. In our preliminary work, we have studied the shrinkage variable selection methods, such as the Lasso [32], adaptive Lasso [33], SCAD [34], which are often used in high-dimensional problems. We found that these methods tend to provide much higher number of false positives when applied to the backcross experiments. We report some simulation results in this regard. It has been shown [33] that the Lasso is not consistent for model selection while the adaptive Lasso is. Therefore, our comparison focuses on the latter. We consider the adaptive Lasso with the tuning parameter chosen either by the 10-fold cross validation (ADA.LASSO.1) or by the BIC (ADA.LASSO.2). For SCAD, we consider the method with the tuning parameter chosen by the with BIC criterion, available in the function `GLMvanISISscad()` of SIS package in R ([35]; SCAD.1), or the method with the function `ncvreg()` in the same-named package in R ([36]; SCAD.2). Tables 5–8 show that the shrinkage methods perform relatively poorly according to criterion (i) comparing to the fence method (see Tables 1–4). The shrinkage methods also have much higher mean false positives than the fence method. The main problem is that the shrinkage methods tend to overfit, leading a fairly good performance with criterion (ii), yet resulting in the inclusion of too many false discoveries. Also note that the performance of the shrinkage methods is much worse in the more complex situation, that is, when the QTL are on the middle of flanking markers. This is shown in Tables 7 & 8, where the result of the criterion (i) is only single digit in percentage, some even 0% in most cases. Similar observations are made in terms of the false discoveries. Overall, the results suggest that the shrinkage methods may not be a good tool for QTL detection in the backcross experiments.

Nevertheless, from our simulation studies (see Tables 5–8), the results of the true discoveries [criterion (ii)] suggest the shrinkage methods appear to be potentially useful for dimensional

reduction since they detect almost all the true QTLs along with possibly extraneous markers, or false discoveries. In our future work, we will consider incorporating such shrinkage methods with the fence method in order to find the parsimonious model in more complex situations when the interaction terms are taken into account.

Acknowledgments

This work was made possible with support from the Oregon Clinical and Translation Research Institute (OCTRI), grant #UL1 RR024140 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. The authors are grateful to a referee for his/her thoughtful comments that led to the improvement of the manuscript.

Appendix: Some derivations and calculations related to Subsection 2.1

We begin with $P(x_j = 1|x_{j-1} = 1) = P(x_j = 0|x_{j-1} = 0) = 1 - \theta$ and $P(x_j = 1|x_{j-1} = 0) = P(x_j = 0|x_{j-1} = 1) = \theta$. Thus, $P(x_j = 1|x_{j-1}) = (1 - \theta)x_{j-1} + \theta(1 - x_{j-1})$. Therefore, we have $E(x_j|x_{j-1}) = 1 \times P(x_j = 1|x_{j-1}) + 0 \times P(x_j = 0|x_{j-1}) = P(x_j = 1|x_{j-1}) = (1 - \theta)x_{j-1} + \theta(1 - x_{j-1}) = \theta + (1 - 2\theta)x_{j-1}$. It follows that $E(x_j|x_{j-k}) = E\{E(x_j|x_{j-1}, \dots, x_{j-k})|x_{j-k}\} = E\{E(x_j|x_{j-1})|x_{j-k}\}$ (property of markov chain process) $= E\{\theta + (1 - 2\theta)x_{j-1}|x_{j-k}\} = \theta + (1 - 2\theta)E(x_{j-1}|x_{j-k}) = \theta + (1 - 2\theta)\{\theta + (1 - 2\theta)E(x_{j-2}|x_{j-k})\} = \theta + \theta(1 - 2\theta) + (1 - 2\theta)^2E(x_{j-2}|x_{j-k}) = \theta\{1 + (1 - 2\theta)\} + (1 - 2\theta)^2\{\theta + (1 - 2\theta)E(x_{j-3}|x_{j-k})\} = \theta\{1 + (1 - 2\theta) + (1 - 2\theta)^2\} + (1 - 2\theta)^3E(x_{j-3}|x_{j-k}) = \dots = \theta\{1 + (1 - 2\theta) + \dots + (1 - 2\theta)^{l-1}\} + (1 - 2\theta)^lE(x_{j-l}|x_{j-k})$.

Thus, we have

$$\begin{aligned} E(x_j|x_{j-k}) &= \theta \left\{ \sum_{j=0}^{k-1} (1 - 2\theta)^j \right\} + (1 - 2\theta)^k x_{j-k}, \quad (l=k) \\ &= \theta \left\{ \frac{1 - (1 - 2\theta)^k}{1 - (1 - 2\theta)} \right\} + (1 - 2\theta)^k x_{j-k} \\ &= \frac{1 - (1 - 2\theta)^k}{2} + (1 - 2\theta)^k x_{j-k}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} x_{j-k}E(x_j|x_{j-k}) &= (1/2) \left\{ 1 - (1 - 2\theta)^k \right\} x_{j-k} + (1 - 2\theta)^k x_{j-k}^2 = (1/2) \left\{ 1 - (1 - 2\theta)^k \right\} x_{j-k} \\ &\quad + (1 - 2\theta)^k x_{j-k} = (1/2) \left\{ 1 + (1 - 2\theta)^k \right\} x_{j-k}. \end{aligned}$$

It follows that $E(x_j x_{j-k}) = E\{x_{j-k}E(x_j|x_{j-k})\} = (1/2)\{1 + (1 - 2\theta)^k\}E(x_{j-k}) = (1/4)\{1 + (1 - 2\theta)^k\}$, and $\text{cov}(x_j, x_{j-k}) = E(x_j, x_{j-k}) - E(x_j)E(x_{j-k}) = (1/4)\{1 + (1 - 2\theta)^k\} - 1/4 = (1 - 2\theta)^k/4$. Therefore, if $s < t$, we have $\text{cov}(x_s, x_t) = \text{cov}\{x_t, x_{t-(t-s)}\} = (1 - 2\theta)^{t-s}/4$; and $\text{var}(x_s) = E(x_s^2) - \{E(x_s)\}^2 = E(x_s) - \{E(x_s)\}^2 = 1/4$. Thus, the contributions to the genetic variance each chromosome are as follow:

$$\begin{aligned}
 \text{var} \left(\sum_{s \in S} \beta_s x_s \right) &= \sum_{s \in S} \beta_s^2 \text{var}(x_s) + \sum_{s \neq t} \beta_s \beta_t \text{cov}(x_s, x_t) \\
 &= \sum_{s \in S} \beta_s^2 \text{var}(x_s) + \sum_{s < t} \beta_s \beta_t \text{cov}(x_s, x_t) + \sum_{s > t} \beta_s \beta_t \text{cov}(x_s, x_t) \\
 &= \sum_{s \in S} \beta_s^2 \text{var}(x_s) + 2 \sum_{s < t} \beta_s \beta_t \text{cov}(x_s, x_t) \\
 &= \frac{1}{4} \sum_{s \in S} \beta_s^2 + 2 \sum_{s < t} \beta_s \beta_t \frac{(1-2\theta)^{t-s}}{4} \\
 &= \frac{1}{4} \sum_{s \in S} \beta_s^2 + \frac{1}{2} \sum_{s < t} \beta_s \beta_t (1-2\theta)^{t-s}.
 \end{aligned}$$

Note that Heritability $\equiv h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$, where σ_g^2 is the variance due to genetic effects. Thus, in the case of 7QTL, we have the following results:

1. On the 1st chromosome, 2 QTLs on 4th and 8th (*coupling* link): $(1/4)(.76^2 + .76^2) + (1/2)(.76^2)(1-2\theta)^4$.
2. On the 2nd chromosome, 2 QTLs on 4th and 8th (*repulsion* link): $(1/4)(.76^2 + .76^2) - (1/2)(.76^2)(1-2\theta)^4$.
3. On the 3rd chromosome, 1 QTL: $(1/4)(.76^2)$.
4. On the 4th chromosome, 1 QTL: same as 3.
5. On the 5th chromosome, 1 QTL: same as 3.

Thus, we have

$$\begin{aligned}
 \sigma_g^2 &= \frac{(.76^2 + .76^2)}{4} + \frac{(.76^2)(1-2\theta)^4}{2} + \frac{(.76^2 + .76^2)}{4} \\
 &\quad - \frac{(.76^2)(1-2\theta)^4}{2} + \frac{3}{4}(.76^2) \approx 1.
 \end{aligned}$$

Also note that $\sigma_e^2 = 1$. Thus, $h^2 \approx .5$. Similarly, in the 9-QTL case, $\beta = .636$.

References

- [1]. Cuénot L. La loi de Mendel et l'hérédité de la pigmentation chez les souris. Arch. Zoöl. exp. et Gén. 1902; 10(3):xxvii–xxx.
- [2]. Castle WE, Allen GM. The heridity of Albinism. Pro. Am. Acad. Arts and Sci. 1903; 38:602–622.
- [3]. Cuénot L. L'hérédité de la pigmentation chez les souris (3), I. Les formules héréditaires. Arch. Zoöl. exp. et Gén. 1904; II L.
- [4]. Cuénot L. Les races pures et leurs combinaisons chez les souris (4). Arch. Zoöl. exp. et Gén. 1905; III L.
- [5]. Castle WE, Little CC. On a modified Mendelian ratio among yellow mice. Science. 1910; 32:868870.
- [6]. Lander ES, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 1989; 121:185–199. [PubMed: 2563713]
- [7]. Zeng Z-B. Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. Proc. Nat. Acad. Sci. 1993; 90:10972–10976. [PubMed: 8248199]
- [8]. Zeng Z-B. Precision mapping of quantitative trait loci. Genetics. 1994; 136:1457–1468. [PubMed: 8013918]
- [9]. Jansen RC. Interval mapping of multiple quantitative trait loci. Genetics. 1993; 135:205–211. [PubMed: 8224820]

- [10]. Jansen RC, Stam P. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*. 1994; 136:1447–1455. [PubMed: 8013917]
- [11]. Akaike, H. Information theory as an extension of the maximum likelihood principle. In: Petrov, BN.; Csaki, F., editors. *Second International Symposium on Information Theory*. Akademiai Kiado; Budapest: 1973. p. 267-281.
- [12]. Schwarz G. Estimating the dimension of a model. *Ann. Statist.* 1978; 6:461–464.
- [13]. Broman, KW. Ph. D. diss. University of California; Berkeley, CA: 1997. Identifying quantitative trait loci in experimental crosses.
- [14]. Broman KW, Speed TP. A model selection approach for the identification of quantitative trait loci in experimental crosses. *J. Roy. Statist. Soc. Ser. B.* 2002; 64:641–656.
- [15]. Broman, KW.; Sen, . A Guide to QTL Mapping with R/ql. Springer; New York: 2009.
- [16]. Ball R. Bayesian methods for quantitative trait loci mapping based on model selection: approximate analysis using the Bayesian information criterion. *Genetics*. 2001; 159:1351–1364. [PubMed: 11729175]
- [17]. Nakamichi R, Ukai Y, Kishino H. Detection of closely linked multiple quantitative trait loci using genetic algorithm. *Genetics*. 2001; 158:463–475. [PubMed: 11333253]
- [18]. Piepho H-P, Gauch HG. Marker pair selection for mapping quantitative trait loci. *Genetics*. 2001; 157:433–444. [PubMed: 11139523]
- [19]. Silanpää MJ, Corander J. Model choice in gene mapping: what and why. *Trends in Genetics*. 2002; 18:301–307. [PubMed: 12044359]
- [20]. Bogdan M, Ghosh JK, Doerge RW. Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci. *Genetics*. 2004; 167:989–999. [PubMed: 15238547]
- [21]. Baierl A, Bogdan M, Frommlet F, Futschik A. On locating multiple interacting quantitative trait loci in intercross design. *Genetics*. 2006; 173:1693–1703. [PubMed: 16624924]
- [22]. Jiang, J. *Linear and Generalized Linear Mixed Models and Their Applications*. Springer; New York: 2007.
- [23]. Jiang J, Rao JS, Gu Z, Nguyen T. Fence methods for mixed model selection. *Ann. Statist.* 2008; 36:1669–1692.
- [24]. Jiang J, Nguyen T, Rao JS. A simplified adaptive fence procedure. *Statist. & Probab. Letters*. 2009; 79:625–629.
- [25]. Nguyen T, Jiang J. Restricted fence method for covariate selection in longitudinal data analysis. *Biostatistics*. 2012; 2(13):303–314. [PubMed: 22180413]
- [26]. Siegmund, D.; Yakir, B. *The statistics of Gene Mapping*. Springer; New York: 2007.
- [27]. Lynch, M.; Walsh, B. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc.; 1998.
- [28]. Fijneman RJA, de Vries SS, Jansen RC, Demant P. Complex interactions of new quantitative trait loci, Sluc1, Sluc2, Sluc3, and Sluc4, that influence the susceptibility to lung cancer in the mouse. *Nature Genetics*. 1996; 14:465–467. [PubMed: 8944028]
- [29]. Fijneman RJA, Jansen RC, Van der Valk MA, Demant P. High frequency of interactions between lung cancer susceptibility genes in the mouse: mapping of Sluc5 to Sluc14. *Cancer Research*. 1998; 58:4794–4798. [PubMed: 9809979]
- [30]. Carlborg Ö, Andersson L, Kinghorn B. The use of a genetic algorithm for simultaneous mapping often multiple interacting quantitative trait loci. *Genetics*. 2000; 155:2003–2010. [PubMed: 10924492]
- [31]. Bogdan, M.; Doerge, RW. Mapping multiple interacting quantitative trait loci with multidimensional genome searches. Department of Statistics, Purdue University; West Lafayette, IN: Technical Report 04-03
- [32]. Tibshirani RJ. Regression shrinkage and selection via the Lasso. *J. Roy. Statist. Soc. Ser. B.* 1996; 16:385–395.
- [33]. Zou H. The adaptive Lasso and its oracle properties. *J. Amer. Statist. Assoc.* 2006; 101:1418–1429.

- [34]. Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* 2001; 96:1348–1360.
- [35]. Fan J, Lv J. Sure independence screening for ultra-high dimensional feature space (with discussion). *J. Roy. Statist. Soc. Ser. B.* 2008; 36:849–911.
- [36]. Breheny P, Huang J. Coordinate descent algorithms for nonconvex penalized regression methods, with applications to biological feature selection. *Ann. Appl. Statist.* 2011; 5:232–253.

Figure 1.
Plots from step 1 of the restricted fence procedure for the first simulation replicate. Nine p^* vs. c_n plots correspond to the nine first halves of the 9 chromosomes

Figure 2. Plots from step 1 of the restricted fence procedure for the first simulation replicate. Nine p^* vs. c_n plots correspond to the nine second halves of the 9 chromosomes

Figure 3.

Plots from step 2 of the restricted fence procedure under the 7-QTL model. Upper left: p^* vs. c_n of the first simulation replicate with sample size $n = 500$. Upper right: same as upper left with smoothing curve. Lower left: p^* vs. c_n of one simulation replicate with sample size $n = 500$. Lower right: p^* vs. c_n of one simulation replicate with sample size of $n = 750$.

Table 1

The 7-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
FB.BIC $\delta-1$	500	%TP M(SD)TP M(SD)FP	81 6.89(.34) 22(.48)	90 7(0) 11(.34)	90 7(0) 11(.34)
FB.BIC $\delta-2$	500	%TP M(SD)TP M(SD)FP	84 6.89(.34) 19(.46)	93 7(0) 08(.3)	93 7(0) 08(.3)
Fence	500	%TP M(SD)TP M(SD)FP	82 6.82(.38) 16(.36)	94 6.94(.23) 04(.3)	94 6.94(.23) 04(.3)
FB.BIC $\delta-1$	750	%TP M(SD)TP M(SD)FP	89 7(0) 12(.35)	89 7(0) 12(.35)	89 7(0) 12(.35)
FB.BIC $\delta-2$	750	%TP M(SD)TP M(SD)FP	94 7(0) 07(.29)	94 7(0) 07(.29)	94 7(0) 07(.29)
Fence	750	%TP M(SD)TP M(SD)FP	98 6.98(.38) 02(.14)	100 7(0) 0(0)	100 7(0) 0(0)

Heritability 50%. E - marker detected as QTL is a true QTL; B.1 - marker detected as QTL is within 10 *cM* from a true QTL; B.2 - marker detected as QTL is within 20 *cM* from a true QTL. In case more than one markers are detected under B.1 or B.2, one is considered as a QTL; the other one(s) as false positive QTL(s). FB.BIC δ - the Forward/Backward BIC δ procedure. Fence - Restricted Fence procedure. FB.BIC $\delta-1$: δ is obtained using the 97.5th percentile of LOD scores. FB.BIC $\delta-2$: δ is obtained using the 95th percentile of LOD scores. %TP - % detecting the true model; M(SD)TP - mean(sd) #'s True positive; M(SD)FP - mean(sd) #'s False positive. *n* - sample size.

Table 2

The 9-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
FB.BIC $\delta-1$	750	%TP M(SD)TP M(SD)FP	83 8.93(0.25) 18(.41)	90 9(0) 11(.34)	90 9(0) 11(.34)
FB.BIC $\delta-2$	750	%TP M(SD)TP M(SD)FP	87 8.93(.25) 14(.37)	94 9(0) 07(.29)	94 9(0) 07(.29)
Fence	750	%TP M(SD)TP M(SD)FP	85 8.85(.35) 11(.31)	93 8.93(.25) 03(.17)	94 8.94(.23) 03(.14)
FB.BIC $\delta-1$	1000	%TP M(SD)TP M(SD)FP	95 9(0) 05(.21)	95 9(0) 05(.21)	95 9(0) 05(.21)
FB.BIC $\delta-2$	1000	%TP M(SD)TP M(SD)FP	98 9(0) 02(.14)	98 9(0) 02(.14)	98 9(0) 02(.14)
Fence	1000	%TP M(SD)TP M(SD)FP	99 8.99(.1) 01(.1)	100 9(0) 0(0)	100 9(0) 0(0)

Heritability 50%. Notations same as in Table 1.

Table 3

The 7-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
FB.BIC δ -1	500	%TP M(SD)TP M(SD)FP	54 6.66(.58) 7(.91)	79 6.98(.14) 29(.55)	88 7(0) 23(.52)
FB.BIC δ -2	500	%TP M(SD)TP M(SD)FP	59 6.63(.61) 57(.79)	87 6.97(.17) 23(.48)	93 7(0) 2(.47)
Fence	500	%TP M(SD)TP M(SD)FP	53 6.4(.7) 56(.8)	78 6.77(.5) 19(.5)	85 6.85(.43) 11(.44)
FB.BIC δ -1	750	%TP M(SD)TP M(SD)FP	44 6.93(.25) 77(.81)	76 7(0) 43(.63)	84 7(0) 37(.56)
FB.BIC δ -2	750	%TP M(SD)TP M(SD)FP	53 (.32) 6(.73)	80 7(0) 33(.56)	86 7(0) 79(.51)
Fence	750	%TP M(SD)TP M(SD)FP	76 6.75(.5) 27(.5)	95 6.97(.17) 05(.21)	97 6.99(.1) 03(.17)

Heritability 40%. E - marker detected is either one of the flanking markers of a true QTL; B.1 - marker detected is within 10 *cM* from either one of the flanking markers of a true QTL; B.2 - marker detected is within 20 *cM* from either one of the flanking markers of a true QTL. In case that more than two markers are detected under B.1 or B.2, two of them are considered as true positives; the other one(s) as false positives. Other notations same as in Table 1.

Table 4

The 9-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
FB.BIC δ - 1	750	%TP M(SD)TP M(SD)FP	52 8.75(.5) 7(.89)	74 8.96(.19) 34(.6)	82 9(0) 26(.48)
FB.BIC δ - 2	750	%TP M(SD)TP M(SD)FP	60 8.73(.52) 53(.75)	83 8.96(.19) 2(.42)	90 8.99(.1) 13(.33)
Fence	750	%TP M(SD)TP M(SD)FP	58 8.47(.71) 47(.68)	84 8.84(.36) 1(.3)	88 8.88(.32) 06(.23)
FB.BIC δ - 1	1000	%TP M(SD)TP M(SD)FP	46 8.9(.38) 75(.9)	75 9(0) 4(.63)	89 9(0) 28(.51)
FB.BIC δ - 2	1000	%TP M(SD)TP M(SD)FP	59 8.9(.38) 53(.78)	87 9(0) 26(.5)	93 9(0) 21(.43)
Fence	1000	%TP M(SD)TP M(SD)FP	75 8.7(.57) 3(.65)	94 8.95(.21) 05(.32)	95 8.97(.17) 03(.22)

Heritability 40%. Notations same as in Table 3.

Table 5

The 7-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
ADA.LASSO.1	500	%TP M(SD)TP M(SD)FP	21 6.97(.17) 2.47(2.53)	31 7(0) 1.68(1.81)	41 7(0) 1.36(1.72)
ADA.LASSO.2	500	%TP M(SD)TP M(SD)FP	18 6.94(.23) 2.03(1.74)	25 7(0) 1.37(1.19)	34 7(0) 1.16(1.13)
SCAD.1	500	%TP M(SD)TP M(SD)FP	11 6.95(.22) 5.98(4.61)	22 7(0) 4.23(3.66)	29 7(0) 3.21(3.02)
SCAD.2	500	%TP M(SD)TP M(SD)FP	43 6.44(1.12) 34.73(42.28)	53 6.49(1.07) 29.29(35.91)	54 6.49(1.07) 24.35(29.84)
ADA.LASSO.1	750	%TP M(SD)TP M(SD)FP	36 7(0) 1.39(1.60)	54 7(0) 92(1.3)	62 7(0) 71(1.12)
ADA.LASSO.2	750	%TP M(SD)TP M(SD)FP	45 7(0) 93(1.17)	51 7(0) 78(.98)	55 7(0) 66(.92)
SCAD.1	750	%TP M(SD)TP M(SD)FP	17 7(0) 4.11(3.65)	25 7(0) 2.9(2.85)	29 7(0) 2.19(2.25)
SCAD.2	750	%TP M(SD)TP M(SD)FP	75 6.73(.91) 11.31(29.01)	81 6.75(.85) 9.55(24.65)	83 6.75(.85) 7.93(20.5)

Heritability 50%. ADA.LASSO.1 - the adaptive lasso procedure with 10-fold cross validation approach. ADA.LASSO.2 - the adaptive lasso procedure with BIC criterion. Other notations same as in Table 1.

Table 6

The 9-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
ADA.LASSO.1	750	%TP M(SD)TP	21 9(0) 2.34(2.21)	32 9(0) 1.61(1.79)	40 9(0) 1.23(1.49)
ADA.LASSO.2	750	%TP M(SD)TP M(SD)FP	24 8.98(.14) 1.66(1.56)	38 9(0) 1.20(1.29)	44 9(0) 98(1.17)
SCAD.1	750	%TP M(SD)TP M(SD)FP	15 9(0) 4.22(4.02)	26 9(0) 2.93(3.1)	33 9(0) 2.15(2.41)
SCAD.2	750	%TP M(SD)TP M(SD)FP	44 8.72(.71) 33.15(41.19)	55 8.75(.7) 26.39(33.07)	58 8.75(.7) 20.4(25.6)
ADA.LASSO.1	1000	%TP M(SD)TP M(SD)FP	28 9(0) 2.22(2.41)	45 9(0) 1.53(1.97)	49 9(0) 1.21(1.62)
ADA.LASSO.2	1000	%TP M(SD)TP M(SD)FP	39 9(0) 1.25(1.53)	46 9(0) 96(1.25)	53 9(0) 78(1.1)
SCAD.1	1000	%TP M(SD)TP M(SD)FP	18 9(0) 3.54(3.3)	26 9(0) 2.47(2.55)	34 9(0) 1.91(2.05)
SCAD.2	1000	%TP M(SD)TP M(SD)FP	69 8.9(.43) 20.93(36.31)	70 8.9(.43) 16.89(29.3)	72 8.9(.43) 13.48(23.47)

Heritability 50%. Notations same as in Table 5.

Table 7

The 7-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
ADA.LASSO.1	500	%TP M(SD)TP M(SD)FP	0 6.92(.31) 9.5(4.71)	1 7(0) 5.59(3.90)	6 7(0) 4.43(3.38)
ADA.LASSO.2	500	%TP M(SD)TP M(SD)FP	0 6.62(.6) 6.32(2.5)	2 6.96(.19) 3.71(1.94)	7 6.98(.14) 2.75(1.61)
SCAD.1	500	%TP M(SD)TP M(SD)FP	0 6.89(.31) 9.24(1.93)	1 7(0) 5.6(1.87)	3 7(0) 4.35(1.68)
SCAD.2	500	%TP M(SD)TP M(SD)FP	4 6.5(1.45) 66.76(27.92)	13 6.53(1.35) 55.42(23.48)	13 6.53(1.35) 47.48(20.12)
ADA.LASSO.1	750	%TP M(SD)TP M(SD)FP	0 7(0) 9.92(5.06)	2 7(0) 5.78(4.10)	6 7(0) 4.29(3.67)
ADA.LASSO.2	750	%TP M(SD)TP M(SD)FP	2 6.86(.37) 5.63(2.58)	8 7(0) 2.70(1.78)	12 7(0) 2.19(1.5)
SCAD.1	750	%TP M(SD)TP M(SD)FP	0 6.97(.17) 11.59(4.21)	0 7(0) 6.9(3.59)	3 7(0) 5.53(3.12)
SCAD.2	750	%TP M(SD)TP M(SD)FP	1 6.52(1.59) 69.08(25.31)	12 6.58(1.38) 57.29(21.49)	12 6.58(.38) 49.09(18.42)

Heritability 40%. E - marker detected is either one of the flanking markers of a true QTL; B.1 - marker detected is within 10 *cM* from either one of the flanking markers of a true QTL; B.2 - marker detected is within 20 *cM* from either one of the flanking markers of a true QTL. In case that more than two markers are detected under B.1 or B.2, two of them are considered as true positives; the other one(s) as false positives. Other notations same as in Table 5.

Table 8

The 9-QTL Case

Procedure	<i>n</i>	Summary	E	B.1	B.2
ADA.LASSO.1	750	%TP M(SD)TP M(SD)FP	0 8.98(.14) 10.2(4.6)	2 9(0) 5.73(3.55)	8 9(0) 4.15(2.98)
ADA.LASSO.2	750	%TP M(SD)TP M(SD)FP	1 8.64(.62) 6.18(2.84)	5 8.94(.23)	12 8.97(.17) 2.1(1.5)
SCAD.1	750	%TP M(SD)TP M(SD)FP	0 8.89(.34) 10.25(3.56)	3 9(0) 5.95(2.92)	7 9(0) 4.18(2.49)
SCAD.2	750	%TP M(SD)TP M(SD)FP	0 8.91(.67) 71.11(15.87)	2 8.92(.58) 55.87(12.66)	3 8.92(.58) 45.17(10.47)
ADA.LASSO.1	1000	%TP M(SD)TP M(SD)FP	0 9(0) 11.36(4.35)	1 9(0) 6.05(3.61)	6 9(0) 4.43(3.05)
ADA.LASSO.2	1000	%TP M(SD)TP M(SD)FP	0 8.85(.35) 5.93(2.31)	8 8.99(.1) 2.59(1.77)	26 9(0) 1.64(1.44)
SCAD.1	1000	%TP M(SD)TP M(SD)FP	0 8.95(.22) 11.04(4.42)	1 9(0) 6.15(3.65)	5 9(0) 4.26(2.93)
SCAD.2	1000	%TP M(SD)TP M(SD)FP	1 8.93(.53) 72.43(12.62)	2 8.93(.53) 56.66(10.39)	3 8.93(.53) 45.85(8.4)

Heritability 40%. Notations same as in Table 7.