ORIGINAL ARTICLE

www.nature.com/hdy

nj

A Fisher scoring algorithm for the weighted regression method of QTL mapping

L Han and S Xu

Department of Botany and Plant Science, University of California, Riverside, CA, USA

An improved weighted least square (LS) method for quantitative trait loci (QTL) mapping using the estimating equation (EE) algorithm was developed recently. The method is more efficient than both the LS and the weighted LS methods and slightly less efficient than the mixture model maximum likelihood (ML) method. The iteration process of the EE algorithm is implicit. We developed a Fisher-scoring algorithm for the weighted LS method. The iteration process is explicit and easy to program. In addition, the method automatically provides an approximate variance–covariance matrix for the estimated QTL parameters as a by-product of the iteration process. As a consequence, a W-test statistic can be used for testing the significance of QTL. To compare the Fisher scoring algorithm with the expectation maximization (EM)-based ML method, we also developed a slightly simplified method to compute the variance–covariance matrix of the estimated parameters under the EM algorithm. *Heredity* (2008) **101,** 453–464; doi:10.1038/hdy.2008.78; published online 13 August 2008

Keywords: QTL mapping; weighted least squares; Fisher scoring

Introduction

The mixture model maximum likelihood (ML) method for quantitative trait loci (QTL) mapping (Lander and Botstein, 1989) is the most efficient method for interval mapping (IM) (Kao, 2000). Least square (Haley and Knott, 1992) and weighted least square (LS; Xu, 1998a, b) methods are approximations of the ML method but with improved computational speed. Recently, Feenstra et al. (2006) developed an improved weighted regression method by extending the simple regression method of Haley and Knott (1992) and the weighted regression method of Xu (1998a, b). The authors made a simple assumption that conditional on marker information, the phenotypic value of individual *j* is normally distributed with mean μ_i and variance σ_1^2 . Feenstra *et al.* (2006) adopted an estimating equation (EE) algorithm to solve for the parameters (regression coefficients and residual error variance).

The exact ML method (Lander and Botstein, 1989) takes into consideration the mixture distribution of the phenotypic value conditional on flanking marker information. The mixture distribution occurs because the genotype of a QTL is unknown. The basic assumption is that the residual error has a known distribution, that is, normal. The simple regression method of Haley and Knott (1992) ignores the uncertainty of the QTL genotype and assumes that $\sigma_j^2 = \sigma^2$ for all $j = 1, \dots, n$, where *n* is the sample size. No other assumption is required. The iteratively reweighted LS (IRLS) method of Xu (1998a and b) takes into account the uncertainty of QTL genotype so that σ_j^2 varies across $j = 1, \dots, n$, but ignores the mixture distribution. In addition, when maximizing the objective function, the method of IRLS treats σ_i^2 as a

Correspondence: Professor S Xu, Department of Botany and Plant Science, University of California, Riverside, CA 92521, USA.

E-mail: xu@genetics.ucr.edu Received 7 March 2008; revised 28 June 2008; accepted 11 July 2008; published online 13 August 2008 constant, although σ_j^2 is a function of μ_{ji} , and thus as a function of QTL effects. The parameter values involved in σ_j^2 are replaced by values in the previous iteration. The EE method of Feenstra *et al.* (2006) takes into account the fact that σ_j^2 is a function of μ_j and maximizes the objective function with respect to parameters occurring in every place of the objective function. Feenstra *et al.* (2006) compared all the methods and showed that ML > EE > IRLS > LS, that is, ML is the most efficient method and LS is the least efficient one.

Feenstra *et al.* (2006) used the EE algorithm to solve for the parameters, but provided no explicit iterative equation. We found that an explicit expression of the iteration exists by using a Fisher scoring algorithm. The iteration equation appears to be simple and thus easy to program. In addition, the method automatically provides a variance–covariance matrix for the estimated QTL effects. This covariance matrix is required to construct a W-test statistic. This test statistic may replace the likelihood-ratio test statistic with a computational advantage over the latter—in that only a single likelihood function (under the full model) is needed.

Methods

Model

The model follows Xu (1998b) and Feenstra *et al.* (2006), which is

$$y_j = U_j \beta + e_j \tag{1}$$

where y_j is the phenotypic value of individual j ($j = 1, \dots, n$), U_j is the expectation of variables indicating the QTL genotype given marker information, β is a vector of QTL effects (including the population mean) and e_j is the residual error. The residual error has mean zero and variance

$$\operatorname{var}(e_i) = \beta^T \Sigma_i \beta + \sigma^2 \tag{2}$$

where Σ_j is the conditional variance–covariance matrix of the QTL indicator variables given marker

information (it is not a summation sign). Note that if the QTL genotype is observed (i.e., the QTL overlaps with a fully informative marker), Σ_j would be zero and thus the residual error variance would be identical to the environmental error variance σ^2 . Definitions and formulations of U_j and Σ_j are given by Xu (1998b) and also described later in this study for an F_2 population. The parameter vector is $\theta = \{\beta, \sigma^2\}$ and the data include y_j, U_j and Σ_i . Let

$$\mu_j = E(y_j) = U_j \beta \tag{3}$$

and

$$\sigma_j^2 = \operatorname{var}(y_j) = \beta^T \Sigma_j \beta + \sigma^2 \tag{4}$$

The likelihood function is based on the assumption of independent $y_i \sim N(\mu_i, \sigma_i^2)$ for $j = 1, \dots, n$. Therefore, the logarithm of the likelihood function is

$$L(\theta) = \sum_{j=1}^{n} L_j(\theta)$$
(5)

where

$$L_{j}(\theta) = -\frac{1}{2}\ln(\sigma_{j}^{2}) - \frac{1}{2\sigma_{j}^{2}}(y_{j} - \mu_{j})^{2}$$
(6)

Fisher scoring algorithm

The partial derivatives of the likelihood for individual *i* with respect to the parameters are

$$\begin{cases} \frac{\partial}{\partial\beta}L_{j}(\theta) = -\frac{1}{2\sigma_{j}^{2}}\left(\frac{\partial}{\partial\beta}\sigma_{j}^{2}\right) + \frac{1}{2}\left[\frac{(y_{j}-\mu_{j})^{2}}{\sigma_{j}^{4}}\left(\frac{\partial}{\partial\beta}\sigma_{j}^{2}\right) + \frac{2(y_{j}-\mu_{j})}{\sigma_{j}^{2}}\left(\frac{\partial}{\partial\beta}\mu_{j}\right)\right] \\ \frac{\partial}{\partial\sigma^{2}}L_{j}(\theta) = -\frac{1}{2\sigma_{j}^{2}}\left(\frac{\partial}{\partial\sigma^{2}}\sigma_{j}^{2}\right) + \frac{1}{2}\left[\frac{(y_{j}-\mu_{j})^{2}}{\sigma_{j}^{4}}\left(\frac{\partial}{\partial\sigma^{2}}\sigma_{j}^{2}\right) + \frac{2(y_{j}-\mu_{j})}{\sigma_{j}^{2}}\left(\frac{\partial}{\partial\sigma^{2}}\mu_{j}\right)\right] \end{cases}$$
(7)

where $\frac{\partial}{\partial \beta} \mu_j = U_j^T$, $\frac{\partial}{\partial \sigma^2} \mu_j = 0$, $\frac{\partial}{\partial \beta} \sigma_j^2 = 2\Sigma_j \beta$ and $\frac{\partial}{\partial \sigma^2} \sigma_j^2 = 1$. Further manipulation on Equation (7) leads to

$$\begin{cases} \frac{\partial}{\partial\beta}L_{j}(\theta) = -\frac{1}{\sigma_{j}^{2}}\Sigma_{j}\beta + \frac{1}{\sigma_{j}^{4}}(y_{j}-\mu_{j})^{2}\Sigma_{j}\beta + \frac{1}{\sigma_{j}^{2}}U_{j}^{T}(y_{j}-\mu_{j})\\ \frac{\partial}{\partial\sigma^{2}}L_{j}(\theta) = -\frac{1}{2\sigma_{j}^{2}} + \frac{1}{2\sigma_{j}^{4}}(y_{j}-\mu_{j})^{2} \end{cases}$$

$$\tag{8}$$

Let $S(\theta) = [\sum_{j=1}^{n} \frac{\partial L_j(\theta)}{\partial \theta}]$ be the score vector and $H(\theta) = [\sum_{j=1}^{n} \frac{\partial^2 L_j(\theta)}{\partial \theta}]$ be the Hessian matrix. The information matrix is $I(\theta) = -E[H(\theta)]$. Using the following identity (Wedderburn, 1974)

$$I(\theta) = -E\left[\sum_{j=1}^{n} \frac{\partial^2 L_j(\theta)}{\partial \theta \ \partial \theta^T}\right] = E\left[\sum_{j=1}^{n} \frac{\partial L_j(\theta)}{\partial \theta} \ \frac{\partial L_j(\theta)}{\partial \theta^T}\right]$$
(9)

we obtained

$$I(\theta) = \begin{bmatrix} \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right) & \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \\ \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right) & \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \end{bmatrix}$$
(10)
$$= \begin{bmatrix} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} U_{j}^{T} U_{j} + 2 \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta \beta^{T} \Sigma_{j} & \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta \\ & \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \beta^{T} \Sigma_{j} & \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \end{bmatrix}$$

The score vector is

$$S(\theta) = \begin{bmatrix} \sum_{j=1}^{n} \frac{\partial L_{j}(\theta)}{\partial \beta} \\ \sum_{j=1}^{n} \frac{\partial^{2} L_{j}(\theta)}{\partial \sigma^{2}} \end{bmatrix}$$
$$= \begin{bmatrix} -\sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta + \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta + \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j}) \\ -\frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} + \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \end{bmatrix}$$
(11)

The derivations of equations (10) and (11) are presented in Appendix A for advanced readers. The information matrix and the score vector are required to perform the following Fisher scoring algorithm for the ML estimation

$$\theta^{(t+1)} = \theta^{(t)} + I^{-1}(\theta^{(t)})S(\theta^{(t)})$$
(12)

where $\theta^{(t)}$ is the parameter value in the *t*-th iteration. Starting with an initial value $\theta^{(0)}$, we iterate equation (12) until a certain criterion of convergence is reached, i.e., $\theta^{(t+1)} \approx \theta^{(t)}$. The ML solution is $\hat{\theta} = \theta^{(t)}$ for *t* satisfying the convergence criterion. The variance-covariance matrix of the estimated parameters is given by

$$\operatorname{var}(\hat{\theta}) \approx I^{-1}(\hat{\theta})$$
 (13)

W-test statistic

When the variance-covariance matrix of the estimated parameters is known, the likelihood ratio tests are not needed. Instead, a W-test statistic can be used (Wald, 1941). We now use an F_2 mating design as an example to show how to construct the W-test statistic. Let

$$X_{j1} = \begin{cases} +1 \text{ for } A_1 A_1 \\ 0 \text{ for } A_1 A_2 \\ -1 \text{ for } A_2 A_2 \end{cases}$$

and (14)

$$X_{j2} = \begin{cases} -1 & \text{for } A_1 A_1 \\ 1 & \text{for } A_1 A_2 \\ -1 & \text{for } A_2 A_2 \end{cases}$$

be indicator variables for the QTL genotype, the linear model of the phenotypic value of individual *j* is

$$y_j = \beta_0 + X_{j1}\beta_1 + X_{j2}\beta_2 + \varepsilon_j \tag{15}$$

where $\beta_0 = \mu$ is the population mean (intercept), $\beta_1 = a$ is the additive effect and $\beta_2 = d$ is the dominance effect and $\varepsilon_j \sim N(0, \sigma^2)$ is the environmental error. Using marker information, we calculate

$$U_j = [1, E(X_{j1}), E(X_{j2})]$$
 (16)

where $E(X_{j1}) = p_j(1) - p_j(-1)$ and $E(X_{j2}) = p_j(0) - [p_j(1) + p_j(-1)]$. The Σ_j , matrix is

$$\Sigma_{j} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \operatorname{var}(X_{j1}) & \operatorname{cov}(X_{j1}, X_{j2}) \\ 0 & \operatorname{cov}(X_{j2}, X_{j1}) & \operatorname{var}(X_{j2}) \end{bmatrix}$$
(17)

where

$$\operatorname{var}(X_{j1}) = p_j(1) + p_j(-1) - [p_j(1) - p_j(-1)]^2,$$
$$\operatorname{var}(X_{j2}) = 1 - \left\{ p_j(0) - [p_j(1) + p_j(-1)] \right\}^2$$

and

$$\operatorname{cov}(X_{j1}, X_{j2}) = [p_j(-1) - p_j(1)] - [p_j(1) - p_j(-1)] \\ \times \{p_j(0) - [p_j(1) + p_j(-1)]\}$$

The probabilities of QTL genotypes conditional on marker information are $p_j(1) = \Pr(X_{j1} = 1 | \text{marker}), p_j(0) = \Pr(X_{j1} = 0 | \text{marker})$ and $p_j(-1) = \Pr(X_{j1} = -1 | \text{marker})$, respectively, for the three genotypes. They are calculated based on the multipoint method of Jiang and Zeng (1997).

The parameter vector is $\theta = \{\beta_0 \ \beta_1 \ \beta_2 \ \sigma^2\}$. Therefore, $\operatorname{var}(\theta)$ is a 4 × 4 matrix. Let $b = [\hat{\beta}_1 \ \hat{\beta}_2]^T$ and

$$\operatorname{var}(b) = \begin{bmatrix} \operatorname{var}(\hat{\beta}_1) & \operatorname{cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \operatorname{cov}(\hat{\beta}_2, \hat{\beta}_1) & \operatorname{var}(\hat{\beta}_2) \end{bmatrix}$$
(18)

be a subset of matrix $var(\hat{\theta})$. The W-test statistic for hypothesis $H_0: \beta_1 = \beta_2 = 0$ is

$$W = b^T [\operatorname{var}(b)]^{-1} b \tag{19}$$

Under H₀, this test statistic follows approximately a χ^2 distribution with two degrees of freedom. Therefore, the W-test statistic is comparable to the likelihood ratio test statistic. The W-test statistic has a simple relationship with the F-test statistic, that is, W = 2F, where the F-test statistic follows an F distribution with a numerator degrees of freedom 2 and a denominator degrees of freedom n-3 under H_0 .

Information matrix of the EM algorithm

The mixture model-based ML method implemented through the expectation maximization (EM) algorithm (Lander and Botstein, 1989) does not have a simple method for calculating the variance–covariance matrix of the estimated parameters. Kao and Zeng's (1997) method for calculating the variance–covariance matrix is quite complicated. If the QTL position is fixed at a particular genome location, that is, the QTL position is not a parameter, their formulas may be simplified. The simplified version of the variance–covariance matrix is relatively easy to program. To compare the Fisher scoring algorithm with the mixture model-based ML method for the variance–covariance matrix of the estimated parameters, we introduced the variance–covariance matrix under the EM algorithm here. The EM algorithm is derived based on the following full data QTL model (assuming that QTL genotypes are observable)

$$y_j = X_{j0}\beta_0 + X_{j1}\beta_1 + X_{j2}\beta_2 + e_j = X_j\beta + e_j$$
(20)

The score function when X_i is fully observed is

$$S(\theta|X) = \begin{bmatrix} \sum_{j=1}^{n} \frac{\partial L_{j}(\theta)}{\partial \beta} \\ \sum_{j=1}^{n} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sigma^{2}} \sum_{j=1}^{n} X_{j}^{T} (y_{j} - X_{j}\beta) \\ -\frac{n}{2\sigma^{2}} + \frac{1}{2\sigma^{4}} \sum_{j=1}^{n} (y_{j} - X_{j}\beta)^{2} \end{bmatrix}$$
(21)

The Hessian matrix (second partial derivatives) is

$$H(\theta|X) = \begin{bmatrix} \sum_{j=1}^{n} \frac{\delta^{2} L_{j}(\theta)}{\partial \beta \partial \sigma^{T}} & \sum_{j=1}^{n} \frac{\delta^{2} L_{j}(\theta)}{\partial \beta \partial \sigma^{2}} \\ \sum_{j=1}^{n} \frac{\delta^{2} L_{j}(\theta)}{\partial \sigma^{2} \partial \beta^{T}} & \sum_{j=1}^{n} \frac{\delta^{2} L_{j}(\theta)}{\partial \sigma^{2} \partial \sigma^{2}} \end{bmatrix}$$
$$= \begin{bmatrix} -\frac{1}{\sigma^{2}} \sum_{j=1}^{n} X_{j}^{T} X_{j} & -\frac{1}{\sigma^{4}} \sum_{j=1}^{n} X_{j}^{T} (y_{j} - X_{j}\beta) \\ -\frac{1}{\sigma^{4}} \sum_{j=1}^{n} (y_{j} - X_{j}\beta) X_{j} & \frac{n}{2\sigma^{4}} - \frac{1}{\sigma^{6}} \sum_{j=1}^{n} (y_{j} - X_{j}\beta)^{2} \end{bmatrix}$$
(22)

The Louis' (1982) information matrix is

$$I(\theta) = -E[H(\theta|X)] - E[S(\theta|X)S^{T}(\theta|X)]$$
(23)

The expectations are taken with respect to the missing value *X* using the posterior probabilities of QTL genotypes. Calculation of the first term $E[H(\theta|X)]$ is straightforward, but the second term $E[S(\theta|X)S^T(\theta|X)]$ is hard to compute. The method of Kao and Zeng (1997) for calculating $E[S(\theta|X)S^T(\theta|X)]$ is complicated. Luo *et al.* (2003) used Monte Carlo simulation to approximate $E[S(\theta|X)S^T(\theta|X)]$, but the method is computationally demanding. We realized that

$$E[S(\theta|X)S^{T}(\theta|X)] = \operatorname{var}[S(\theta|X)]$$
(24)

because $E[S(\theta|X)] = 0$ when $\theta = \hat{\theta}$, where $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ . The variance–covariance matrix of the scores is

$$\operatorname{var}[S(\theta|X)] = \begin{bmatrix} \sum_{j=1}^{n} \operatorname{var}\left(\frac{\partial L_{j}(\theta)}{\partial \beta}\right) & \sum_{j=1}^{n} \operatorname{cov}\left(\frac{\partial L_{j}(\theta)}{\partial \beta}, \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \\ \sum_{j=1}^{n} \operatorname{cov}\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}, \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right) & \sum_{j=1}^{n} \operatorname{var}\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \end{bmatrix}$$
(25)

which is not difficult to calculate. Note that the expectations and the variance–covariance matrices are calculated with respect to the missing value *X* using the posterior probabilities of QTL genotypes (conditional on both marker and phenotype information and the

456

estimated parameter value). Therefore, the information matrix for the EM estimated QTL parameters is

$$I(\theta) = -E[H(\theta|X)] - E[S(\theta|X)S^{T}(\theta|X)]$$

= -E[H(\theta|X)] - var[S(\theta|X)] (26)

because $E[S(\theta|X)S^T(\theta|X)] = var[S(\theta|X)]$ as shown in equation 24. The estimated variance–covariance matrix for $\hat{\theta}$ is $var(\hat{\theta}) = I^{-1}(\hat{\theta})$. This variance–covariance matrix will be compared to that obtained from the weighted LS method.

Results

Monte carlo simulation

The purpose of the simulation study is to verify the Fisher scoring algorithm. The final result of the Fisher scoring algorithm is identical to the result of EE (Feenstra *et al.*, 2006), because both algorithms maximize the same likelihood function. Extensive simulation studies for the improved weighted regression method have been performed by Feenstra *et al.* (2006). Therefore, we only evaluated the Fisher scoring algorithm under one situation. We placed a QTL in the middle of a 10 cM marker interval. The two markers were fully informative. The simulated parameter values were

$\theta =$	$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \sigma^2 \end{bmatrix}$	=	$\begin{bmatrix} \mu \\ a \end{bmatrix}$	=	10.0 0.400 0.125 1.000	
	$\lfloor \sigma^2 \rfloor$		σ^2		1.000	

The additive and dominance effects of the QTL explained 7.3% and 1.43% of the phenotypic variance, respectively. Overall, the QTL contributed 8.73% of the phenotypic variance. The sample size of the simulated F_2 population was n = 300. The simulation experiment was replicated 200 times. The QTL position was fixed at the true location (5 cM away from either flanking marker) and only θ was estimated.

For comparison purpose, we also analyzed the data using the simple regression or the LS method of Haley and Knott (1992), the IRLS method of Xu (1998a, b) and the mixture model ML method of Lander and Botstein (1989). The Fisher scoring method developed in this study is denoted by FISHER. The EE algorithm is simply a different algorithm from the Fisher scoring algorithm for the same problem. Both EE and FISHER maximize the same likelihood function, and thus both generate the same ML estimates of the parameters. For the three methods with iterations, the same initial values of parameters were used, which are

$$\left[\mu^{(0)} \ a^{(0)} \ d^{(0)} \ \sigma^{2(0)}\right] = \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}$$

The average estimates of the parameters and their standard deviations obtained from the 200 replicated simulations are listed in Table 1. The estimated parameters for all the four methods are very close to the true parameters. This verified all methods, including the Fisher scoring method developed here. The new method took about five iterations to converge, whereas ML and IRLS took about seven and four iterations, respectively, to converge to the same criterion $(||\theta^{(t+1)} - \theta^{(t)}|| \le 10^{-8})$. The computing time for the Fisher scoring method took about 0.1 s per replication, which is approximately the same as the time for the IRLS method. Both FISHER and IRLS are faster than ML.

Special algorithm is required to obtain the variancecovariance matrix of the estimated QTL parameters for the ML method (Kao and Zeng, 1997; Luo et al., 2003). The Fisher scoring method, however, provides such a covariance matrix as a by-product of the iteration process. To evaluate the accuracy of the estimated variance-covariance matrix, we compared the 'predicted' covariance matrix with the 'realized' covariance matrix. The predicted covariance matrix was obtained as follows. For the kth replicated simulation $(k = 1, \dots, 2001)$ we calculated $V^{(k)} = \operatorname{var}(\hat{\theta})$ using equation (13) and then took the average $\bar{V} = \frac{1}{200} \sum_{k=1}^{200} V^{(k)}$ as the 'predicted' covariance matrix. The realized covariance matrix was calculated using the following approach. Let $\hat{\theta}^{(k)}$ be the estimated parameters from the *k*th replicated simulation ($k = 1, \dots, 200$). The 'realized' covariance matrix was defined as

$$V = \frac{1}{200} \sum_{k=1}^{200} (\hat{\theta}^{(k)} - \bar{\hat{\theta}}) (\hat{\theta}^{(k)} - \bar{\hat{\theta}})^T$$

where $\overline{\hat{\theta}} = \frac{1}{200} \sum_{k=1}^{200} \hat{\theta}^{(k)}$. The predicted and realized covariance matrices obtained from 200 replicated simulations are given in Table 2. Clearly, the predicted covariance matrix is very close to the realized covariance matrix.

Numerical evaluation

A dataset from an F_2 mouse population consisting of 110 individuals was used as an example for demonstration.

 Table 1
 Mean estimates of parameters and the standard deviations (in parentheses) of the estimated parameters obtained from 200 replicated simulations

Parameter	True value	thod			
		LS	ML	IRLS	FISHER
μ	10.000	9.999 (0.060)	9.999 (0.060)	9.999 (0.060)	9.999 (0.059)
Â	0.400	0.418 (0.093)	0.418 (0.092)	0.418 (0.093)	0.418 (0.091)
D	0.125	0.119 (0.065)	0.119 (0.066)	0.118 (0.065)	0.119 (0.066)
σ^2	1.000	1.012 (0.080)	0.992 (0.079)	1.002 (0.080)	0.992 (0.079)
Iteration	_	<u> </u>	7.135	3.885	5.085
Time	_	0.0817	0.1526	0.1095	0.1005

Abbreviations: IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

The average number of iterations and the average computing time (s) per replicate are given in the last two rows.

The data were published by Lan *et al.* (2006) and are freely available from the internet (see Lan *et al.*, 2006 for the website address). A preliminary analysis showed that there was a QTL for the trait of 10th week body weight on the second chromosome between markers D2Mit194 and D2Mit263 (result not given). The putative position of the QTL is at 95.7 cM, whereas the two flanking markers are located at 85.4 and 98.7 cM, respectively. QTL parameters were estimated assuming that the position of the QTL is known (fixed at 95.7 cM).

The iteration process of the Fisher scoring algorithm is given in Table 3. It took seven iterations to converge when $\left[\mu^{(0)} a^{(0)} d^{(0)} \sigma^{2(0)}\right] = [1 \ 0 \ 0 \ 1]$ was used as the initial values of the parameters. If the LS estimates of QTL parameters were used as the initial values, only five iterations were required to converge to the same criterion, $||\theta^{(t+1)} - \theta^{(t)}|| \leq 10^{-8}$ (data not shown).

The data were also analyzed using the other three methods (LS, IRLS and ML) and the results are given in Table 4. The estimated parameters from LS and IRLS are more alike to each other than to the parameters estimated from ML and FISHER, which are almost identical to each other. The W-test statistics for all the four methods (LS, IRLS, ML and FISHER) were compared with the likelihood-ratio test statistics and they are indeed similar to each others (see last two rows of Table 4).

The variance–covariance matrices for the estimated parameters for the four methods are given in Table 5. The covariance matrices for the four methods are very similar. To further validate the accuracy of the covariance matrix of the Fisher scoring method, we performed a bootstrap analysis (Efron, 1979) with 1000 replicated

Table 2 Comparison of the predicted with the realized variancecovariance matrices of the estimated parameters for the Fisher scoring algorithm

		Predicted						
	μ̂	â	â	$\hat{\sigma}^2$				
$\hat{\mu} \\ \hat{a} \\ \hat{d} \\ \hat{\sigma}^2$	0.00338 -0.00004 -0.00001 0.00000	-0.00004 0.00767 -0.00007 -0.00037	-0.00001 -0.00007 0.00436 -0.00021	0.00000 -0.00037 -0.00021 0.00683				
		Rea	lized					
	μ̂	â	â	$\hat{\sigma}^2$				
$\hat{\mu} \\ \hat{a} \\ \hat{d} \\ \hat{\sigma}^2$	0.00354 -0.00015 -0.00014 -0.00015	-0.00015 0.00836 0.00016 0.00001	-0.00014 0.00016 0.00436 0.00023	-0.00015 0.00001 0.00023 0.00618				

samples to draw an empirical covariance matrix for the estimated parameters. The bootstrapped covariance matrix for the FISHER method is also presented in Table 5. We can see that the bootstrapped covariance matrix is similar to the predicted one from the Fisher scoring method, although some relatively large deviations have occurred for some covariance elements.

Interval mapping and composite interval mapping (CIM) This section presents the results of IM and CIM (Zeng 1994) for the same mouse data (Lan et al., 2006). The mouse genome has 19 chromosomes (excluding the sex chromosome). The data investigated contain $110 F_2$ mice and 193 markers covering about 1800 cM of the entire genome. The trait of interest is still the 10th week body weight. The entire genome was scanned with 1 cM increment using the IM approach (under the single QTL model). The log of odds (LOD) score profiles (converted from the likelihood ratio test statistic profiles) for all the four methods (LS, IRLS, ML and FISHER) are given in Figure 1. The four methods generate almost identical result (the profiles overlap). Permutation tests showed that the critical value for the LOD score was 2.985, close to 3 (data not shown). Therefore, we used LOD 3 as the approximate critical value for controlling the genome wise Type I error of 0.05 for declaration of statistical significance. Three QTL passed the LOD 3 criterion for all the four methods. The estimated positions and effects for the three QTL are given in Table 6. The three QTL detected explain 10-20% of the phenotypic variance each. Again, the four methods give almost identical result for the estimated QTL parameters.

We have identified three QTL using the IM approach. Since IM uses a single QTL model, the estimated QTL parameters are subject to bias (Zeng, 1994). We now adopt the CIM approach to handle multiple QTL. We scanned the entire genome again but used the markers nearby the three identified QTL (in the IM) as cofactors to control the background noise. The LOD score profiles are presented in Figure 2. Using the same LOD 3 as the approximate critical value, we detected two QTL, one remains in chromosome 2 and the other one occurs in chromosome 13. Again, the four methods are almost indistinguishable. The estimated QTL parameters are given in Table 7. The two detected QTL individually explain a less proportion of the phenotypic variance compared with the results of the IM, which is expected for the CIM approach.

Figure 3 represents the LOD score profiles for the FISHER method developed in this study under theIM and the CIM frameworks. Again, the IM is a single QTL

Table 3 Iteration process of the Fisher scoring algorithm of QTL mapping for the mouse data

Iteration	μ	а	d	σ^2	$ \theta^{(t+1)} - \theta^{(t)} $
0	1	0	0	1	_
1	60.002269	-2.892456	0.2439982	3549.8115	3149388.3
2	60.002134	-2.893374	0.2453587	31.071243	3095383.3
3	59.99052	-2.975879	0.3830536	30.879357	0.0156806
4	59.988608	-2.989226	0.3904464	30.822373	0.0008709
5	59.988479	-2.990227	0.3913964	30.819513	2.5261E-6
6	59.988465	-2.990331	0.3914711	30.819225	2.4867E-8
7	59.988464	-2.99034	0.3914783	30.819201	1.792E-10

Table 4 Estimated QTL parameters and their standard errors (in parentheses) for the mouse data from four different methods

Parameter	Method					
	LS	ML	IRLS	FISHER		
μ	60.002 (0.543)	59.991 (0.539)	60.002 (0.548)	59.988 (0.539)		
a	-2.892 (0.827)	-2.990 (0.826)	-2.905 (0.831)	-2.990 (0.814)		
d	0.244 (0.587)	0.384 (0.590)	0.256 (0.590)	0.391 (0.579)		
σ^2	31.700 (4.274)	30.797 (4.226)	31.849 (4.295)	30.819 (4.204)		
Number of iteration	_	8	5	7		
Likelihood ratio	11.706	12.149	11.736	11.613		
W-test	12.239	13.096	12.227	13.553		

Abbreviations: IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

 Table 5 Variance-covariance matrix of the estimated QTL parameters for the mouse data

Method	$\hat{\mu}$	â	â	$\hat{\sigma}^2$
LS				
$\hat{\mu}$	0.295	0.067	-0.019	0.000
â	0.067	0.684	-0.058	0.000
â	-0.019	-0.058	0.345	0.000
$\hat{\sigma}^2$	0.000	0.000	0.000	18.271
ML				
$\hat{\mu}$	0.290	0.066	-0.020	0.017
â	0.066	0.683	-0.083	0.212
â	-0.020	-0.083	0.348	-0.083
$\hat{\sigma}^2$	0.017	0.212	-0.083	17.859
IRLS				
$\hat{\mu}$	0.300	0.068	-0.019	0.000
â	0.068	0.690	-0.059	0.000
â	-0.019	-0.059	0.348	0.000
$\hat{\sigma}^2$	0.000	0.000	0.000	18.443
FISHER				
$\hat{\mu}$	0.290	0.065	-0.018	0.014
â	0.065	0.663	-0.056	0.143
â	-0.018	-0.056	0.335	0.004
$\hat{\sigma}^2$	0.014	0.143	0.004	17.671
FISHER-b ^a				
$\hat{\mu}$	0.265	-0.017	0.039	-0.017
â	-0.017	0.543	0.035	0.390
â	0.039	0.035	0.332	0.176
$\hat{\sigma}^2$	-0.017	0.390	0.176	17.628

Abbreviations: IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

^aFISHER-b obtained from 1000 bootstrapped replicates of the primary dataset.

model, whereas the CIM is a multiple QTL model. The two different approaches do show significant difference. The CIM is an improved approach and is highly recommended over the IM approach.

Discussion

The Fisher scoring algorithm developed in this study and the EE algorithm developed by Feenstra *et al.* (2006) produce identical result, because both maximize the same likelihood function. Therefore, they are two different algorithms for the same method, called the improved weighted LS method. Properties of the

improved weighted LS method have been investigated thoroughly by Feenstra et al. (2006). Therefore, we only provided results of simulations and real data analysis for a simple situation: the QTL position is known. The purpose of this study is to demonstrate that a Fisher scoring algorithm can be used to estimate QTL parameters. Given that FISHER is identical to EE, why do we bother to develop such an algorithm? Several reasons may justify the new algorithm. One is that Fisher scoring is an important algorithm in genetics and this study gives another example of its application to genetics. Another reason is that the EE algorithm introduced by Feenstra et al. (2006) has no explicit expression of the iteration process, whereas the Fisher scoring algorithm does. More importantly, the EE was performed in two steps within an iteration, the β -step and the σ^2 -step. As a result, variance-covariance matrix of the estimated parameters is not available for the EE algorithm, at least it is not a by-product of the iteration process, but the Fisher scoring algorithm provides such a matrix as a byproduct of the iteration process. We were informed that Feenstra B (personal communication) did derive the information matrix for the parameters under the EE algorithm in his thesis, but the result was not published.

An algorithm of maximization, like the Fisher scoring algorithm, that provides an easy way for calculating the variance–covariance matrix of the estimated parameters may be more preferable than the EM algorithm for the mixture model ML method, because a W-test statistic can be used for significance test of QTL. The W-test statistic is similar to the likelihood-ratio test, but it is computationally less intensive than the latter. For the Fisher scoring algorithm, three different test statistics (additive, dominance and both) can be generated with only one objective function (the likelihood function for the full model) for maximization, whereas the EM implemented ML method requires maximization of several likelihood functions (under the full model and various reduced models).

The Fisher scoring implemented weighted LS method has some similarity to the quasilikelihood method (Wedderburn, 1974). Both require only the first (mean) and second (variance) moments of an observed data point to be known functions of parameters and no other assumptions are needed. The Fisher scoring method, however, cannot be replaced by the quasilikelihood method, because the latter requires the variance to be a known function of the mean for each observed data point.



Weighted least squares mapping

Figure 1 Log of odds (LOD) test statistic profiles for the four methods (LS, IRLS, ML and FISHER) using the interval mapping (IM) approach. The 19 chromosomes are merged into a single genome. The dash–dot horizontal line along the top of the graph represents the threshold value of LOD 3. The main and minor tick marks on the horizontal axis separate the chromosomes and indicate the position of markers, respectively. IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

Table 6 Estimated QTL parameters for the mice data from the interval mapping analysis

QTL	Method	LOD	QTL position (cM)	Nearest marker	μ	а	d	σ^2	h^2
QTL 1	LS	3.424	175.257	D2Mit9	58.966	-3.574 (0.848)	1.622 (1.241)	29.989	0.190
	ML	3.469	174.286	D2Mit9	59.244	-3.444(0.780)	1.131 (1.060)	29.664	0.174
	IRLS	3.412	175.257	D2Mit9	59.122	-3.116 (0.792)	1.455 (1.074)	29.694	0.153
	FISHER	3.468	174.286	D2Mit9	59.204	-3.459(0.784)	1.207 (1.059)	29.633	0.176
QTL 2	LS	3.299	206.81	D2Mit274	59.338	-2.820(0.731)	1.659 (1.060)	30.419	0.133
	ML	3.299	206.81	D2Mit274	59.354	-2.803(0.730)	1.626 (1.056)	30.393	0.131
	IRLS	3.299	206.81	D2Mit274	59.351	-2.806 (0.730)	1.630 (1.056)	30.426	0.131
	FISHER	3.299	206.81	D2Mit274	59.350	-2.807(0.730)	1.633 (1.056)	30.390	0.132
QTL 3	LS	3.504	233.06	D2Mit263	59.665	-3.328(0.793)	0.816 (1.166)	30.111	0.159
	ML	3.508	233.06	D2Mit263	59.699	-3.282(0.745)	0.747 (1.049)	29.678	0.157
	IRLS	3.485	234.01	D2Mit263	59.817	-3.024(0.753)	0.566 (1.052)	29.857	0.135
	FISHER	3.508	233.06	D2Mit263	59.704	-3.280 (0.752)	0.736 (1.044)	29.683	0.157

The standard errors of estimated QTL effects (square roots of the variances of estimated QTL effects) are given in parentheses.



Figure 2 Log of odds (LOD) test statistic profiles for the four methods (LS, IRLS, ML and FISHER) using the composite interval mapping (CIM) approach. The 19 chromosomes are merged into a single genome. The dash–dot horizontal line along the top of the graph represents the threshold value of LOD 3. The main and minor tick marks on the horizontal axis separate the chromosomes and indicate the position of markers, respectively. IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

459

Weighted least squares mapping L Han and S Xu

QTL	Method	LOD	QTL position (cM)	Nearest marker	μ	а	d	σ^2	h^2
QTL 1	LS	3.299	206.81	D2Mit274	59.338	-2.820 (0.731)	1.659 (1.060)	30.393	0.133
~	ML	3.299	206.81	D2Mit274	59.354	-2.803(0.730)	1.626 (1.056)	30.426	0.131
	IRLS	3.299	206.81	D2Mit274	59.351	-2.806(0.730)	1.630 (1.056)	30.390	0.131
	FISHER	3.299	206.81	D2Mit274	59.350	-2.807(0.730)	1.633 (1.056)	30.419	0.132
QTL 2	LS	3.454	1270.82	D13Mit66	56.323	2.603 (0.682)	2.143 (1.015)	24.533	0.156
	ML	3.456	1270.82	D13Mit66	56.336	2.562 (0.677)	2.097 (1.008)	24.588	0.151
	IRLS	3.456	1270.82	D13Mit66	56.346	2.562 (0.677)	2.091 (1.007)	24.525	0.151
	FISHER	3.456	1270.82	D13Mit66	56.341	2.567 (0.677)	2.097 (1.007)	24.583	0.152

Table 7 Estimated QTL parameters for the mice data from the composite interval mapping (CIM) analysis

Abbreviations: IRLS, iteratively reweighted LS; LS, least square; ML, maximum likelihood.

The standard errors of estimated QTL effects (square roots of the variances of estimated QTL effects) are given in parentheses.



Figure 3 Log of odds (LOD) test statistic profiles for the Fisher scoring method (FISHER) under the interval mapping (IM) and composite interval mapping (CIM) frameworks. The 19 chromosomes are merged into a single genome. The dash–dot horizontal line along the top of the graph represents the threshold value of LOD 3. The main and minor tick marks on the horizontal axis separate the chromosomes and indicate the position of markers, respectively.

Although the EM implemented ML is not the focus of this study, the simple method for calculating the variance of the score functions, $var[S(\theta|X)]$, is new and cost efficient relatively to the method of Kao and Zeng (1997) and the Monte Carlo method of Luo *et al.* (2003).

Any quantitative traits may be controlled by more than one QTL. The IM approach uses a model that contains only a single QTL. Therefore, the model will never be the correct one if multiple QTL exist. However, people using the IM approach can still detect multiple QTL, simply by evaluating the number of peaks in the test statistic profile that pass the critical value of the test statistic. The estimated QTL effects will be biased if multiple linked QTL exist. There are two ways to handle multiple QTL. One is the multiple IM developed by Kao et al (1999), where a multiple QTL model is fit and the number of QTL included is determined by a model selection algorithm, for example, stepwise regression. The other method is called the CIM where markers with significant effects are selected and then fit into the model to control the background noise (Zeng, 1994). In this study, we modified the Fisher scoring algorithm to fit the CIM model. We first used the IM approach to identify markers. We found three separate peaks in chromosome 2 of the mouse genome. We then selected three markers that are close to the three peaks as cofactors and fit the composite mapping model. Once we fit the CIM model, the three QTL in chromosome 2 became one and an additional QTL in chromosome 13 was detected. Therefore, the CIM does show different result from the IM. We did not fit a true multiple QTL mapping model, because the focus of this study is to develop the Fisher scoring method rather than evaluating the multiple QTL mapping model. The CIM does serve as a simple method to handle multiple QTL. The actual multiple QTL model under the Fisher scoring algorithm is not hard to fit, but requires substantial computational time. Let *q* be number of QTL fit to the model, the multiple QTL model is

$$y_{j} = \mu + \sum_{k=1}^{q} U_{jk} \beta_{k} + e_{j}$$
(27)

where U_{jk} is the conditional expectation of the QTL genotype indicator variable for locus *k* and β_k is a vector of QTL effects for the *k*th locus. The residual error has mean zero and variance

$$\operatorname{var}(e_j) = \sum_{k=1}^{q} \beta_j^T \sum_{jk} \beta_k + \sigma^2$$
(28)

where Σ_{jk} is the conditional variance–covariance matrix of the *k*th QTL genotype indicator variable. The score vector and the information matrix are equivalent to those obtained for the single QTL model except that the dimensionalities of the vector and matrix are increased.

Epistatic effects may also contribute to the genetic variance of a quantitative trait. However, they are not as important as the main QTL effects for many quantitative traits of agronomy crops (Xu and Jia, 2007). Extension to models with epistatic effects is tedious and will digress from the main focus of this study. Therefore, we only introduce the main effect model under the Fisher scoring algorithm. Feenstra et al (2006) investigated the properties of the EE algorithm and found out that the EE algorithm can be substantially more efficient that the LS and weighted LS methods when epistatic effects are fit to the model. Since the Fisher scoring algorithm is identical to the EE algorithm in terms of parameter estimation and significance test, we expect that the Fisher scoring algorithm will also show the same advantage as the EE algorithm in handling epistatic effects.

Finally, we developed a new SAS procedure called PROC QTL, which has an option to conduct QTL mapping using the Fisher scoring method. The SAS code is given in Appendix B.

Acknowledgements

This project was supported by the National Research Initiative (NRI) Plant Genome of the USDA Cooperative State Research, Education and Extension Service (CSREES) 2007–02784 to SX.

References

- Efron B (1979). Bootstrap methods: another look at the jackknife. *Ann Statist* 7: 1–26.
- Feenstra B, Skovgaard IM, Broman KW (2006). Mapping quantitative trait loci by an extension of the Haley–Knott regression method using estimating equations. *Genetics* **173**: 2269–2282.
- Haley CS, Knott SA (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315–324.
- Kao CH, Zeng ZB (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 CH, Zeng ZB, Teasdale RD (1999). Multiple interval mapping for quantitative trait loci. *Genetics* 152: 1203–1216.
- Kao CH (2000). On the differences between maximum likelihood and regression interval mapping in the analysis of quantitative trait loci. *Genetics* **156**: 855–865.
- Jiang C, Zeng ZB (1997). Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. *Genetica* 101: 47–58.
- Lander ES, Botstein D (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.
- Lan H, Chen M, Flowers JB, Yandell BS, Stapleton DS, Mata CM *et al.* (2006). Combined expression trait correlations and expression quantitative trait locus mapping. *PLoS Genet* **2**: 51–61.
- Louis TA (1982). Finding the observed information matrix when using the EM algorithm. J Roy Statist Soc Ser B 44: 226–233.
- Luo L, Mao Y, Xu S (2003). Correcting the bias in estimation of genetic variances contributed by individual QTL. *Genetica* 119: 107–113.
- Wedderburn RWM (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika* **61**: 439–447.

- Wald A (1941). Asymptotically most powerful tests of statistical hypotheses. Ann Math Stat 12: 1–19.
- Xu S (1998a). Further investigation on the regression method of mapping quantitative trait loci. *Heredity* 80: 364–373.
- Xu S (1998b). Iteratively reweighted least squares mapping of quantitative trait loci. *Behav Genet* 28: 341–355.
- Xu S, Jia Z (2007). Genomewide analysis of epistatic effects for quantitative traits in barley. *Genetics* 175: 1955–1963.
- Zeng ZB (1994). Precision mapping of quantitative trait loci. *Genetics* **136**: 1457–1468.

Appendix A

Derivation of the score vector and the information matrix The log-likelihood function is the sum of $L_j(\theta)$ across all *n* individuals, where

$$L_{j}(\theta) = -\frac{1}{2}\ln(\sigma_{j}^{2}) - \frac{1}{2\sigma_{j}^{2}}(y_{j} - \mu_{j})^{2}$$
(A1)

The expectation and variance of y_j is $\mu_j = U_j\beta$ and $\sigma_j^2 = \beta^T \Sigma_j \beta + \sigma^2$, respectively. The first partial derivatives of the likelihood for individual *j* with respect to the parameters are

$$\begin{cases} \frac{\partial}{\partial\beta}L_{j}(\theta) = -\frac{1}{2\sigma_{j}^{2}}\left(\frac{\partial}{\partial\beta}\sigma_{j}^{2}\right) + \frac{1}{2}\left[\frac{(y_{j}-\mu_{j})^{2}}{\sigma_{j}^{4}}\left(\frac{\partial}{\partial\beta}\sigma_{j}^{2}\right) + \frac{2(y_{j}-\mu_{j})}{\sigma_{j}^{2}}\left(\frac{\partial}{\partial\beta}\mu_{j}\right)\right] \\ \frac{\partial}{\partial\sigma^{2}}L_{j}(\theta) = -\frac{1}{2\sigma_{j}^{2}}\left(\frac{\partial}{\partial\sigma^{2}}\sigma_{j}^{2}\right) + \frac{1}{2}\left[\frac{(y_{j}-\mu_{j})^{2}}{\sigma_{j}^{4}}\left(\frac{\partial}{\partial\sigma^{2}}\sigma_{j}^{2}\right) + \frac{2(y_{j}-\mu_{j})}{\sigma_{j}^{2}}\left(\frac{\partial}{\partial\sigma^{2}}\mu_{j}\right)\right] \end{cases}$$
(A2)

where $\frac{\partial}{\partial \beta} \mu_j = U_j^T$, $\frac{\partial}{\partial \sigma^2} \mu_j = 0$, $\frac{\partial}{\partial \beta} \sigma_j^2 = 2\Sigma_j \beta$ and $\frac{\partial}{\partial \sigma^2} \sigma_j^2 = 1$. Further manipulation on the above equations leads to

$$\begin{cases} \frac{\partial}{\partial\beta}L_j(\theta) = -\frac{1}{\sigma_j^2}\Sigma_j\beta + \frac{1}{\sigma_j^4}(y_j - \mu_j)^2\Sigma_j\beta + \frac{1}{\sigma_j^2}U_j^T(y_j - \mu_j)\\ \frac{\partial}{\partial\sigma^2}L_j(\theta) = -\frac{1}{2\sigma_j^2} + \frac{1}{2\sigma_j^4}(y_j - \mu_j)^2 \end{cases}$$
(A3)

The score vector simply takes the sum of the above individual wise first partial derivatives across all individuals

$$S(\theta) = \begin{bmatrix} \sum_{j=1}^{n} \frac{\partial L_{j}(\theta)}{\partial \beta} \\ \sum_{j=1}^{n} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \end{bmatrix}$$
$$= \begin{bmatrix} -\sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta + \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta + \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j}) \\ -\frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} + \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \end{bmatrix}$$
(A4)

This concludes the derivation of equation (11) in the main text.

The information matrix is $I(\theta) = -E[H(\theta)]$. Using the following identity (Wedderburn, 1974)

$$I(\theta) = -E\left[\sum_{j=1}^{n} \frac{\partial^2 L_j(\theta)}{\partial \theta \, \partial \theta^T}\right] = \sum_{j=1}^{n} E\left[\frac{\partial L_j(\theta)}{\partial \theta} \frac{\partial L_j(\theta)}{\partial \theta^T}\right]$$
(A5)

we obtained

$$I(\theta) = \begin{bmatrix} \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right) & \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \\ \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right) & \sum_{j=1}^{n} E\left(\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right) \end{bmatrix}$$
(A6)

We now provide the details of each element of the above matrix. The products of the two partial derivatives involved in the information matrix are given as follows.

1. The first product:

$$\begin{split} & \left[\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right] \\ &= \left[-\frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j})\right] \left[-\frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta\right]^{T} \\ &+ \left[-\frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j})\right] \\ &\times \left[\frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta\right]^{T} \\ &+ \left[-\frac{1}{\sigma_{j}^{2}} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{4}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta + \frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j})\right] \\ &\times \left[\frac{1}{\sigma_{j}^{2}} U_{j}^{T} (y_{j} - \mu_{j})\right]^{T} \end{split}$$
(A7)

Further manipulation on the above equation leads to

$$\begin{split} &\left[\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right] \\ = &\left[\frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta \beta^{T} \Sigma_{j} - \frac{1}{\sigma_{j}^{6}} (y_{j} - \mu_{j})^{2} \Sigma_{j} \beta \beta^{T} \Sigma_{j} - \frac{1}{\sigma_{j}^{4}} U_{j}^{T} \beta^{T} \Sigma_{j} (y_{j} - \mu_{j})\right] \\ &+ \left[-\frac{1}{\sigma_{j}^{6}} \Sigma_{j} \beta \beta^{T} \Sigma_{j} (y_{j} - \mu_{j})^{2} + \frac{1}{\sigma_{j}^{8}} (y_{j} - \mu_{j})^{4} \Sigma_{j} \beta \beta^{T} \Sigma_{j} \right. \\ &\left. + \frac{1}{\sigma_{j}^{6}} U_{j}^{T} \beta^{T} \Sigma_{j} (y_{j} - \mu_{j})^{3}\right] \\ &+ \left[-\frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta U_{j} (y_{j} - \mu_{j}) + \frac{1}{\sigma_{j}^{6}} (y_{j} - \mu_{j})^{3} \Sigma_{j} \beta U_{j} \right. \\ &\left. + \frac{1}{\sigma_{j}^{4}} U_{j}^{T} U_{j} (y_{j} - \mu_{j})^{2}\right] \end{split}$$

$$(A8)$$

Taking the expectation value with respect to the observed phenotypic value y_{j} , we obtained

$$E\left[\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right]$$

$$= \left[\frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j} - \frac{1}{\sigma_{j}^{6}}E[(y_{j} - \mu_{j})^{2}]\Sigma_{j}\beta\beta^{T}\Sigma_{j}$$

$$- \frac{1}{\sigma_{j}^{4}}U_{j}^{T}\beta^{T}\Sigma_{j}E(y_{j} - \mu_{j})\right]$$

$$+ \left[-\frac{1}{\sigma_{j}^{6}}\Sigma_{j}\beta\beta^{T}\Sigma_{j}E[(y_{j} - \mu_{j})^{2}] + \frac{1}{\sigma_{j}^{8}}E[(y_{j} - \mu_{j})^{4}]\Sigma_{j}\beta\beta^{T}\Sigma_{j}$$

$$+ \frac{1}{\sigma_{j}^{6}}U_{j}^{T}\beta^{T}\Sigma_{j}E[(y_{j} - \mu_{j})^{3}]\right]$$

$$+ \left[-\frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta U_{j}E(y_{j} - \mu_{j}) + \frac{1}{\sigma_{j}^{6}}E[(y_{j} - \mu_{j})^{3}]\Sigma_{j}\beta U_{j}$$

$$+ \frac{1}{\sigma_{j}^{4}}U_{j}^{T}U_{j}E[(y_{j} - \mu_{j})^{2}]\right]$$
(A9)

Note that

$$E[(y_j - \mu_j)^k = \begin{cases} 0 & \text{for odd } k \\ \sigma^2 & \text{for } k = 2 \\ 3\sigma_j^4 & \text{for } k = 4 \end{cases}$$
(A10)

Therefore, the above equation becomes

$$E\left[\frac{\partial L_{j}(\theta)}{\partial \beta}\frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right] = \left[\frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j} - \frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j}\right] \\ + \left[-\frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j} + \frac{3}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j}\right] + \left[\frac{1}{\sigma_{j}^{2}}U_{j}^{T}U_{j}\right] \\ = \frac{1}{\sigma_{j}^{2}}U_{j}^{T}U_{j} + \frac{2}{\sigma_{j}^{4}}\Sigma_{j}\beta\beta^{T}\Sigma_{j}$$
(A11)

Therefore

$$\sum_{j=1}^{n} E\left[\frac{\partial L_{j}(\theta)}{\partial \beta} \frac{\partial L_{j}(\theta)}{\partial \beta^{T}}\right] = \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} U_{j}^{T} U_{j} + \sum_{j=1}^{n} \frac{2}{\sigma_{j}^{4}} \Sigma_{j} \beta \beta^{T} \Sigma_{j}$$
(A12)

We now have concluded the derivation for the first diagonal block of the information matrix.

2. The second product:

$$\begin{split} & \left[\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}} \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right] \\ &= \left[-\frac{1}{2\sigma_{j}^{2}} + \frac{1}{2\sigma_{j}^{4}}(y_{j} - \mu_{j})^{2}\right] \left[-\frac{1}{2\sigma_{j}^{2}}\right] \\ &+ \left[-\frac{1}{2\sigma_{j}^{2}} + \frac{1}{2\sigma_{j}^{4}}(y_{j} - \mu_{j})^{2}\right] \left[\frac{1}{2\sigma_{j}^{4}}(y_{j} - \mu_{j})^{2}\right] \\ &= \left[\frac{1}{4\sigma_{j}^{4}} - \frac{1}{4\sigma_{j}^{6}}(y_{j} - \mu_{j})^{2}\right] + \left[-\frac{1}{4\sigma_{j}^{6}}(y_{j} - \mu_{j})^{2} + \frac{1}{4\sigma_{j}^{8}}(y_{j} - \mu_{j})^{4}\right] \\ &\qquad (A13) \end{split}$$

The expectation of the above product is

$$E\left[\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right] = \left[\frac{1}{4\sigma_{j}^{4}} - \frac{1}{4\sigma_{j}^{6}}E(y_{j} - \mu_{j})^{2}\right] + \left[-\frac{1}{4\sigma_{j}^{6}}E(y_{j} - \mu_{j})^{2} + \frac{1}{4\sigma_{j}^{8}}E(y_{j} - \mu_{j})^{4}\right] = \left[\frac{1}{4\sigma_{j}^{4}} - \frac{1}{4\sigma_{j}^{4}}\right] + \left[-\frac{1}{4\sigma_{j}^{4}} + \frac{3}{4\sigma_{j}^{4}}\right] = \frac{1}{2\sigma_{j}^{4}}$$
(A14)

Therefore

$$\sum_{j=1}^{n} E\left[\frac{\partial L_j(\theta)}{\partial \sigma^2} \frac{\partial L_j(\theta)}{\partial \sigma^2}\right] = \sum_{j=1}^{n} \frac{1}{2\sigma_j^4}$$
(A15)

We have now completed the derivation for the second diagonal block of the information matrix.

3. The third product:

The expectation of the above equation is

$$E\left[\frac{\partial L_{j}(\theta)}{\partial \beta} \quad \frac{\partial L_{j}(\theta)}{\partial \sigma^{2}}\right]$$

$$=\left[\frac{1}{2\sigma_{j}^{4}}\Sigma_{j}\beta - \frac{1}{2\sigma_{j}^{6}}E(y_{j}-\mu_{j})^{2}\Sigma_{j}\beta - \frac{1}{2\sigma_{j}^{4}}U_{j}^{T}E(y_{j}-\mu_{j})\right]$$

$$+\left[-\frac{1}{2\sigma_{j}^{6}}\Sigma_{j}\beta E(y_{j}-\mu_{j})^{2} + \frac{1}{2\sigma_{j}^{8}}E(y_{j}-\mu_{j})^{4}\Sigma_{j}\beta + \frac{1}{2\sigma_{j}^{6}}U_{j}^{T}E(y_{j}-\mu_{j})^{3}\right]$$

$$=\left[\frac{1}{2\sigma_{j}^{4}}\Sigma_{j}\beta - \frac{1}{2\sigma_{j}^{4}}\Sigma_{j}\beta\right] + \left[-\frac{1}{2\sigma_{j}^{4}}\Sigma_{j}\beta + \frac{3}{2\sigma_{j}^{4}}\Sigma_{j}\beta\right]$$

$$=\frac{1}{\sigma_{j}^{4}}\Sigma_{j}\beta \qquad (A17)$$

Therefore

$$\sum_{j=1}^{n} E \left[\frac{\partial L_j(\theta)}{\partial \beta} \quad \frac{\partial L_j(\theta)}{\partial \sigma^2} \right] = \sum_{j=1}^{n} \frac{1}{\sigma_j^4} \Sigma_j \beta$$
(A18)

We have now concluded the derivation of the offdiagonal block of the information matrix. The information matrix is symmetrical, and thus the other offdiagonal block simply takes the transposition of the above quantity. Substituting all the four elements into the information matrix, we obtained the final expression of the information matrix,

$$I(\theta) = \begin{bmatrix} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{2}} U_{j}^{T} U_{j} + 2 \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta \beta^{T} \Sigma_{j} & \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \Sigma_{j} \beta \\ \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \beta^{T} \Sigma_{j} & \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\sigma_{j}^{4}} \end{bmatrix}$$
(A19)

which is the matrix given in equation 10 of the main text.

Appendix B

The SAS code of QTL mapping

We developed a new SAS procedure called PROC QTL, which has an option to conduct QTL mapping using four different methods (LS, IRLS, ML and FISHER). The option FISHER is the Fisher scoring method developed in this study. The SAS code is given in this appendix. People who want to use this program to conduct QTL mapping must be a regular SAS users. Before running the procedure, users need to ask for a trial version of the executable file from the authors and install the new procedure to the existing SAS software package licensed from the SAS company. We will deliver the software (executable file) along with a user manual for the software. The QTL procedure is not a built-in SAS procedure, but once you install this customized SAS procedure into your system, you can execute the procedure just like you do for any other builtin SAS procedures without noticing the difference between this customized procedure and any other regular SAS procedures. The SAS code is given in the next page.

The SAS code to run PROC QTL:

filename aa 'C:\Users\gen.csv'; filename bb 'C:\Users\phe.csv'; filename cc 'C:\Users\map.csv'; filename dd 'C:\Users\result.csv';

data one;

infile aa dlm = ',' lrecl = **20000**; input id (m1-m193)(\$);

data two;

infile bb dlm = ',' lrecl = 20000; input id sex \$ y;

data three;

infile cc dlm = ',' lrecl = 20000; input marker \$ position chromoso;

data mouse; merge one two;

by id;

PROC QTL data = mouse map = three out = result method = 'fisher' step = 1;

model y = ; matingtype 'f2'; genotype A1A1 = 'A' A1A2 = 'H' A2A2 = 'B'; estimate 'a' = 1 0 - 1 'd' = -1 1 - 1;

run;

data four;

set result; file dd dlm = ','; put trait chr marker position n_iter conv_err LRT Wald ve intercpt a d var_1 cov_1_2 var_2;

run;

q**uit**;