ORIGINAL ARTICLE

Bayesian analysis for genetic architecture of dynamic traits

L Min[1,2], R Yang[1], X Wang[1] and B Wang[1]

[1]*Department of Animal Science, School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai, PR China and* [2]*Department of Animal Science, College of Animal Science and Technology, Qingdao Agricultural University, Qingdao, PR China*

The dissection of the genetic architecture of quantitative traits, including the number and locations of quantitative trait loci (QTL) and their main and epistatic effects, has been an important topic in current QTL mapping. We extend the Bayesian model selection framework for mapping multiple epistatic QTL affecting continuous traits to dynamic traits in experimental crosses. The extension inherits the efficiency of Bayesian model selection and the flexibility of the Legendre polynomial model fitting to the change in genetic and environmental effects with time. We illustrate the proposed method by simultaneously detecting the main and epistatic QTLs for the growth of leaf age in a doubled-haploid population of rice. The behavior and performance of the method are also shown by computer simulation experiments. The results show that our method can more quickly identify interacting QTLs for dynamic traits in the models with many numbers of genetic effects, enhancing our understanding of genetic architecture for dynamic traits. Our proposed method can be treated as a general form of mapping QTL for continuous quantitative traits, being easier to extend to multiple traits and to a single trait with repeat records.

Heredity (2011) **106,** 124–133; doi:10.1038/hdy.2010.20; published online 24 March 2010

Introduction

The process of formation and development of a biological trait may have different temporal and spatial properties. Such a trait whose phenotype changes with time or quantitative factor is known as a dynamic trait. Biologically speaking, the change in phenotype of the trait can be due to different genes that turn on or off at various times. In other words, the dynamic trait is governed by some genes whose genetic effects change with time. Studying the changing laws of these gene effects and their mutual relationships can enhance our understanding of the genetic architecture of dynamic traits.

The genetic mechanism of dynamic traits has been observed in practice on mapping quantitative trait loci (QTL) for some fixed time points in dynamic traits, with separate analysis (Cheverud *et al.*, 1983; Nuzhdin *et al.*, 1997; Verhaegen *et al.*, 1997; Emebiri *et al.*, 1998; Wu *et al.*, 1999), joint analysis (Jiang and Zeng, 1995; Korol *et al.*, 1995; Ronin *et al.*, 1995; Eaves *et al.*, 1996; Knott and Haley, 2000) or conditional analysis (Yan *et al.*, 1998a, b; Wu *et al.*, 2002). Subsequently, the research focus has been gradually shifted toward betterfitting changing laws in genetic effects of QTL genotypes and genes. Wu and his colleagues (Ma *et al.*, 2002; Wu *et al.*, 2004a, b, c; Wu and Lin, 2006) proposed a functional mapping strategy constructed within the context of interval

mapping, where the mean vectors of QTL genotypes within a time interval are modeled by a biologically meaningful mathematical equation, and the covariance matrix is modeled in terms of its time series autocorrelation structure (Ma *et al.*, 2002). Fitting the Legendre polynomials to the time-dependent genetic effects of markers outside the test interval, Yang *et al.* (2007) presented a flexible nonparametric approach for composite functional mapping of dynamic traits. Although these functional mapping strategies have emerged as a powerful tool for mapping dynamic trait loci, using nonlinear biologically meaningful mathematical model-to-model changes of QTL genotype effects may limit their extension to a multiple QTL model. Moreover, there is still a lack of biologically meaningful mathematical models for most dynamic traits.

The Legendre polynomial has been extensively used by animal geneticists and breeders to fit changes in breeding values for milk production and other dynamic traits (Kirkpatrick and Heckman, 1989; Kirkpatrick *et al.*, 1990; Schaeffer, 2004). This has stimulated several usages of the Legendre polynomial in QTL mapping for dynamic traits. For example, Yang *et al.* (2004) and Huang *et al.* (2005) replaced the Logistic curve with the Legendre polynomial and made functional mapping suitable for dynamic traits with an arbitrary shape. In Macgregor *et al.* (2005), the Legendre polynomial was applied to QTL mapping for longitudinal traits in pedigrees. They adopted the traditional random regression model in which the vector of polynomial regression coefficients (genetic effects) for each animal is treated as a random vector sampled from a multivariate normal distribution. For line crosses, Yang *et al.* (2006) proposed an interval mapping method for dynamic traits by using

Correspondence: *Dr R Yang, School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai 200240, PR China.*
E-mail: runqingyang@sjtu.edu.cn

the Legendre polynomial to model the population mean, QTL effects and time-dependent environmental effects. On the basis of this interval mapping method, Yang and Xu (2007) subsequently developed a Bayesian shrinkage analysis framework to simultaneously map genome-wide QTLs with multiple main effects for dynamic traits.

The dissection of the genetic architecture of quantitative traits, including the number and locations of QTLs and their main and epistatic effects, becomes an important topic in current QTL mapping. In fact, the unknown number of QTLs and the possibly huge number of epistatic effects make the issue extremely complex. A promising approach for solving the issue is the Bayesian model selection framework, which has been developed to identify epistatic QTL for regular quantitative traits (Yi *et al.*, 2005, 2007b) and for ordinal traits (Yi *et al.*, 2007a), but not for dynamic traits.

In this study, we will extend the Bayesian model selection for mapping interacting QTLs developed by Yi *et al.* (2005, 2007b) to dynamic traits in experimental crosses. The extension is realized by embedding the Legendre polynomials in QTL effects and taking into account the individual-specific time-dependent random environmental effects in the genetic model. Our extension inherits the efficiency of the Bayesian model selection and the flexibility of the Legendre polynomial model fitting to change in genetic and environmental effects with time, and can fairly quickly identify interacting QTLs for dynamic traits in models with large numbers of genetic effects, which are demonstrated by analyzing the simulated data and real data on leaf age growth in rice.

Methods

Genetic model
We start with a simple population including only two segregating genotypes at each locus, such as a backcross (BC), double-haploid lines or recombinant inbred lines. For mapping the QTL of dynamic traits, phenotypes of repeated measurements in time interval $[t_0, t_m]$ and molecular marker data need to be collected on n individuals. Assume that there are q QTL responsible for changing the trajectory of dynamic traits. The phenotypic value $y_i(t)$ of individual i measured at time t can be then described by the following multiple interacting QTL model (Kao and Zeng, 2002):

$$y_i(t) = \mu(t) + \sum_{j=1}^{q} \gamma_j x_{ij}\beta_j(t) + \sum_{j=1}^{q-1}\sum_{k=j+1}^{q} \gamma_{jk} z_{ijk}\delta_{jk}(t)$$
$$+ \xi_i(t) + \varepsilon_i(t)$$

(1)

where $\mu(t)$ is the population mean at time t; $\beta_j(t)$ for $j = 1, 2,…,q$ is the additive effect of the jth QTL at time point t; $\delta_{jk}(t)$ is the epistatic effect between jth QTL and kth QTL for $j = 1, 2,…,q-1$; $k = j+1, j+2,…,q$; x_{ij} is a genotype indicator variable for individual i at locus j and is defined as 1 for one genotype and -1 for the other genotype; z_{ijk} is the dummy variable for epistatic effect between jth QTL and kth QTL on ith individual, $z_{ijk} = x_{ij}x_{ik}$; $\gamma_\bullet$ is a binary variable for each genetic effect, indicating whether the corresponding effect is included ($\gamma_\bullet = 1$) or excluded ($\gamma_\bullet = 0$) from model (1); $\xi_i(t)$ is an individual-specific time-dependent random environ-

mental effect, distributed as $N(0, \sigma_\xi^2(t))$; and $\varepsilon_i(t)$ is a time-independent random residual error, following the normal distribution with mean 0 and variance σ^2. Notice that by inferring $\gamma_\bullet$, the Bayesian model selection enables the Markov Chain Monte Carlo (MCMC) sampling for QTL parameters to be conducted in a reduced model space (Carlin and Chib, 1995; Yi, 2004).

The Legendre polynomial of p orders is chosen to fit the changing trajectories of the population mean, QTL effects and residual error. Let $\psi(t)$ be the basis of the Legendre polynomial (see Yang *et al.*, 2006) and stipulate that $\mu(t) = \psi(t)\mu$, $\beta_j(t) = \psi(t)\beta_j$, $\delta_{jk}(t) = \psi(t)\delta_{jk}$ and $\xi_i(t) = \psi(t)\xi_i$, where μ, β_j, δ_{jk} and ξ_i are the $p+1$ vectors of the regression coefficients. Model (1) can then be rewritten as

$$y_i(t) = \psi(t)\mu + \sum_{j=1}^{q} \gamma_j x_{ij}\psi(t)\beta_j$$
$$+ \sum_{j=1}^{q-1}\sum_{k=j+1}^{q} \gamma_{jk} z_{ijk}\psi(t)\delta_{jk} + \psi(t)\xi_i + \varepsilon_i(t)$$

(2)

Assume that ξ_i is i.i.d. $N(0, \Sigma)$, where Σ is a $(p+1) \times (p+1)$ positive definite covariance matrix.

For simplicity of description, we assume that each individual has m measurements at m different time points and that the time points are common for all individuals. However, our method can accommodate the data from arbitrary time points. Let $y_i = [y_i(t_0) \ y_i(t_1) … y_i(t_m)]^T$ be a $(m+1) \times 1$ column vector for the repeated measurements of the dynamic traits, and define $\psi = [\psi^T(t_0) \ \psi^T(t_1) … \psi^T(t_m)]$ as a $(p+1) \times (m+1)$ matrix. In matrix notation, model (2) becomes

$$y_i = \psi^T\mu + \sum_{j=1}^{q} \gamma_j x_{ij}\psi^T\beta_j$$
$$+ \sum_{j=1}^{q-1}\sum_{k=j+1}^{q} \gamma_{jk} z_{ijk}\psi^T\delta_{jk} + \psi^T\xi_i + \varepsilon_i$$

(3)

where $\varepsilon_i(t) = [\varepsilon_{it_0} \ \varepsilon_{it_1} … \varepsilon_{it_m}]^T$ is a $(m+1) \times 1$ vector for the environmental errors with $\varepsilon_i \sim N(0, I\sigma^2)$, where I is an $(m+1) \times (m+1)$ identity matrix. The conditional expectation of model (3) given the fixed effects, such as population mean and genetic effects, is

$$E(y_i) = M_i$$
$$= \psi^T\mu + \sum_{j=1}^{q} \gamma_j x_{ij}\psi^T\beta_j + \sum_{j=1}^{q-1}\sum_{k=j+1}^{q} \gamma_{jk} z_{ijk}\psi^T\delta_{jk}$$

(4)

and the variance–covariance matrix is

$$\text{Var}(y_i) = V = \psi^T\Sigma\psi + I\sigma^2$$

(5)

for all $i = 1, 2,…, n$.

Bayesian mapping
Similar to the cases for regular quantitative traits, the Bayesian mapping framework implemented in MCMC algorithms for dynamic traits mainly consists of six consecutive parts: (1) to establish the likelihood function for phenotypes according to the given genetic model reflecting the relationship between phenotypes and unknown parameters; (2) to specify the prior distribution

for each unknown parameter; (3) to form the joint posterior distribution by multiplying the likelihood function from step 1 by all prior distributions from step 2; (4) to obtain the conditional posterior distribution for each unknown parameter by fixing other parameters in joint posterior distribution; (5) to draw MCMC samples for each unknown parameter from the corresponding conditional posterior distributions and (6) to analyze the posterior samples for each parameter and statistically characterize them. In contrast to regular quantitative traits, however, Bayesian mapping for dynamic traits is more complex because of the consideration of time dependence of QTL effects and random environmental effects on traits of interest.

Likelihood function: Denote the phenotypic observations $y = \{y_i\}$ for $i = 1, 2, \ldots, n$, the unknown parameters $\gamma = \{\gamma_j \ \gamma_{jk}\}$, $X = \{x_{ij} \ z_{ijk}\}$, $\lambda = \{\lambda_j\}$ with λ_j being the position of the jth QTL and $\theta = \{\mu \ \beta_j \ \delta_{jk} \ \xi_i \ \Sigma \ \sigma^2\}$ for $j = 1, 2, \ldots, q$; $k = j+1, \ j+2, \ldots, q$. The likelihood function is the conditional distribution of y given γ, X and θ, which is denoted by:

$$p(y|\gamma, X, \theta, \lambda) = \prod_{i=1}^{n} p(y_i|\gamma, X, \theta, \lambda) \propto |V|^{-n/2}$$
$$\times \exp\left[\sum_{i=1}^{n} (y_i - M_i)^{\mathrm{T}} V^{-1} (y_i - M_i)\right]$$

Prior distribution: Notice that the genetic effects of QTLs on dynamic traits in models (1)–(3) are equivalent to nesting the Legendre polynomial within the genetic effects of QTL on regular quantitative traits. Therefore, in Bayesian mapping for dynamic traits, choices of the upper bound L and specification of the prior on γ and λ should be the same as those for regular quantitative traits. As described by Yi *et al*. (2005), we take L as $l_0 + 3\sqrt{l_0}$, where l_0 is the prior expected number of QTLs and is determined according to initial investigations with traditional methods. The binary indicator γ is assumed to have an independent prior $p(\gamma) = \prod w_{\bullet}^{\gamma_{\bullet}} \cdot (1 - w_{\bullet})^{(1-\gamma_{\bullet})}$, where $w_{\bullet}$ is the prior inclusion probability for a certain QTL effect and equals the predetermined hyperparameter w_m for main effects or w_e for epistatic effects, respectively. Priors on λ are assumed to be independent and uniformly distributed over the entire genome, that is, QTL positions have a uniform prior information.

The prior for the population mean μ is $N(\mu_0, \Sigma_0)$. We can empirically set

$$\mu_0 = \bar{b} = \frac{1}{n}\sum_{i=1}^{n} b_i \quad \text{and}$$
$$\Sigma_0 = \frac{1}{n-1}\sum_{i=1}^{n} (b_i - \bar{b})^{\mathrm{T}}(b_i - \bar{b})$$

where $b_i = (\psi^{\mathrm{T}}\psi)^{-1}\psi^{\mathrm{T}}y_i$ and is a vector of regression coefficients obtained by fitting the individual dynamic trajectory.

We propose the following hierarchical mixture prior for each additive genetic effect,

$$\beta_j \sim N_{p+1}(0, \Sigma_j) \quad \text{for} \quad j = 1, 2, \ldots, p$$
with

$$\Sigma_j = \gamma_j c \left(\psi^{\mathrm{T}} V^{-1} \psi \sum_{i=1}^{n} x_{ij}^2\right)^{-1}$$

and c being taken to n such that the prior variance of each fixed effect stays approximately the same as n increases. Similarly, we take the prior distribution for epistatic effect as

$$\delta_{jk} \sim N_{p+1}(0, \Sigma_{jk}) \quad \text{for} \quad j = 1, 2, \ldots, q;$$
$$k = j+1, j+2, \ldots, q$$

with

$$\Sigma_{jk} = \gamma_{jk} c \left(\psi^{\mathrm{T}} V^{-1} \psi \sum_{i=1}^{n} z_{ijk}^2\right)^{-1}$$

The random effects ξ_i are assumed to have an independent multivariate normal distribution, that is, $\xi_i \sim N_{p+1}(0, S_a)$ with the hyperparameter S_a being a $(p+1) \times (p+1)$ matrix.

An inverse Wishart prior is chosen for the covariance matrix of regression coefficients for random environmental effect, denoted by $\Sigma \sim \mathrm{IW}(v_a, v_a S_a)$ with v_a being a hyperparameter.

The residual variance is assigned to be a scaled inverse χ^2 distribution, that is, $\sigma_0^2 \sim \mathrm{IC}(v_e, \frac{1}{v_e s_e})$ with v_e and s_e being hyperparameters.

Genotypes of missing markers were generated randomly in each iteration on the basis of the probability inferred jointly from the nearest nonmissing flanking markers and the phenotype. The probability from the missing marker locus is treated as the prior probability. After incorporation of the marker (Locus) effects through the phenotype, the probability becomes the posterior probability, which is used to generate the missing marker genotype from multinomial distribution. The detailed calculation of posterior probabilities for missing marker genotypes can be found in Wang *et al*. (2005).

The joint prior of all parameters takes the product of the priors of individual parameters.

MCMC algorithm: In general, the joint posterior density derived from likelihood function and the joint priors of all parameters are intractable analytically. However, MCMC methods such as the Gibbs sampler (Gelman *et al*., 1995) and the Metropolis–Hastings algorithm (Metropolis *et al*., 1953; Hastings, 1970) can be used to draw samples, from which features of marginal distributions of interest can be inferred.

Within the framework of the Bayesian Model selection, the upper bound L on the number of QTLs is not only given, but also the released sampling value for $\gamma_{\bullet}$ at current iteration determined which genetic effect and QTL position will be drawn or estimated at the next iteration. This allowed us to conduct Bayesian sampling for QTL parameters in a reasonably reduced model space, thus greatly decreasing the computational demand.

On the basis of marginal posterior distribution for each parameter (shown in Appendix A), we implement MCMC sampling by the following computationally efficient process:

(1) Evenly partition the entire genome into small intervals (1 or 2 cM long) by a number of points and restrict putative QTLs to these fixed points. Estimate all expected values of indicator variables X for putative QTL by using conditional probabilities of their genotypes on two flanking markers.

(2) Divide the entire genome into L equal intervals and put one QTL in the middle of each interval.

(3) Initialize all variables with some legal values or values sampled from their prior distributions;

(4) Update the population mean μ;

(5) Update the binary indicators γ with an efficient Metropolis–Hastings algorithm (Kohn *et al.*, 2001; Yi *et al.*, 2007a);

(6) Update the additive QTL effects β_j corresponding to $\gamma_j = 1$;

(7) Update the epistatic QTL effects δ_{jk} corresponding to $\gamma_{jk} = 1$;

(8) Update the residual variance σ^2;

(9) Update the QTL position λ_j on those fixed points, corresponding to $\gamma_\bullet = 1$;

(10) Repeat steps (4)–(9) until the Markov chain reaches a desirable length.

As the order of V equals the number of repeat measurements for dynamic traits, it is hard to calculate the inverse and determinant for V when there are a large number of repeat measurements. In practices of MCMC sampling, therefore, the inverse and determinant for V need to be solved in the form of the reduced dimension. The detailed derivation of the simplified formula is given in Appendix B.

For analyzing the models with multiple interacting QTLs and only multiple main-effect QTLs by using the Bayesian model selection, we write the program to implement MCMC sampling in Matlab, which can be available from the authors on request.

Post-MCMC analysis: The posterior sample can be used to infer the genetic architecture of quantitative traits, including the number and locations of QTL and their main and epistatic effects. Before doing these, we need to monitor the mixing behavior and convergence rates of MCMC algorithms by visually inspecting trace plots of the sample values of scalar quantities of interest or by using formal diagnostic methods provided in the package R/coda (Plummer *et al.*, 2006). Model averaging accounts for model uncertainty provide more robust inference compared with a single optimal model approach (Raftery *et al.*, 1997; Ball, 2001; Sillanpää and Corander, 2002) and are therefore used to assess the characteristics of genetic architecture by averaging over possible models weighted by their posterior probabilities. We can use various methods to graphically and numerically summarize and interpret the posterior samples. The posterior inclusion probability for each locus is estimated as its frequency in the posterior samples; taking the prior probability into consideration, we use Bayes factors (BFs) to show evidence for inclusion against exclusion of each QTL effect. The BF for a locus or QTL effect is defined as the ratio of the posterior odds to the prior odds for inclusion against exclusion of the QTL locus or effect (Kass and Raftery, 1995). Generally, a threshold of BF is taken to 3 or

$2 \ln BF = 2.1$, for declaring statistical significance for each QTL effect (Kass and Raftery, 1995).

Real data analysis

A doubled-haploid (DH) population with 111 lines was generated by crossing an indica rice variety Gui-630 and a japonica rice variety Taiwanjing. A linkage map composed of 175 RFLP markers was constructed using the DH population, covering a total length of 1225 cM with average spacing of 7 cM (Weng *et al.*, 2000). This DH population was grown with replicates in a field trial (Zhou *et al.*, 2001). For each plant, the number of developed leaves on the main stem was counted, and the length of the developing leaf was measured every 3–7 days from day 30 after sowing until the full development of the leaf. These measured data were used to estimate the leaf age of a plant (y) using

$$y = \text{Number of developed leaves}$$
$$+ \frac{\text{Length of the developing leaf}}{\text{Final length of the developing leaf}}$$

The time points of measurements counted by the numbers of days after the seeds were $t = (5\ 8\ 13\ 18\ 21\ 26\ 32\ 39)$.

We select the Legendre polynomial of order 2 to model changes of population mean and genetic effects with growth time on the basis of the changing law of phenotypes of trait. The data are analyzed by adopting the maximum likelihood method (Yang *et al.*, 2006) and Bayesian method, respectively.

Before Bayesian sampling, we partitioned each chromosome with a 1-cM grid, which resulted in 1214 possible loci across the genome. The actual values for the hyper parameters are $S_a = S_e = 0.5I$, $v_a = p + 1$ and $v_e = 0$. The initial values of all variables were sampled from their prior distributions. For all Bayesian analyses, the MCMC sampling ran for 200 000 cycles after discarding the first 2000 burn-ins. The chain was thinned by recording one sample in every 40 samples, yielding 5000 samples for posterior Bayesian analysis.

With interval mapping based on maximum likelihood (Yang *et al.*, 2006), the five significant QTLs were detected on chromosomes 1, 5, 9, 10 and 12, respectively. Under the nonepistatic analysis, the number of significant QTLs detected in interval mapping was taken as the prior number of main-effect QTLs, and the upper bound of the number of QTLs was then calculated as $L = 5 + 3\sqrt{5} = 12$. The graph of the BFs is displayed as the bottom plot in Figure 1. It can be seen that besides five QTLs identified by interval mapping, four more clear peaks arise on chromosomes 2, 3, 4 and 7. Moreover, all relative BFs of the nine peaks found above are greater than the significant threshold of 3.

The epistatic analysis also took the expected number of main-effect QTLs to 5, as nonepistatic analysis did, and the expected number of all QTLs was chosen as 8. The maximum number of QTLs was then $L = 8 + 3\sqrt{8} = 16$.

The estimated population mean and covariance matrix for random regression coefficients for the individual-specific environmental effects are

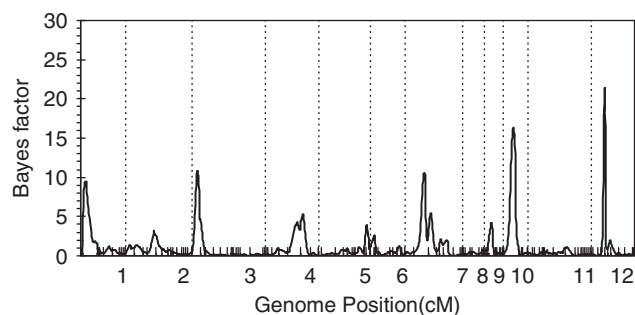$$\hat{\mu} = \begin{bmatrix} 11.28 \\ 2.78 \\ -0.41 \end{bmatrix}$$

Figure 1 Profiles of Bayes factor with Bayesian nonepistatic analysis. Chromosomes are separated by the vertical dotted lines and marker positions are indicated by the ticks on the horizontal axis.

Figure 2 Profiles of Bayes factor with Bayesian epistatic analysis. Chromosomes are separated by the vertical dotted lines and marker positions are indicated by the ticks on the horizontal axis.

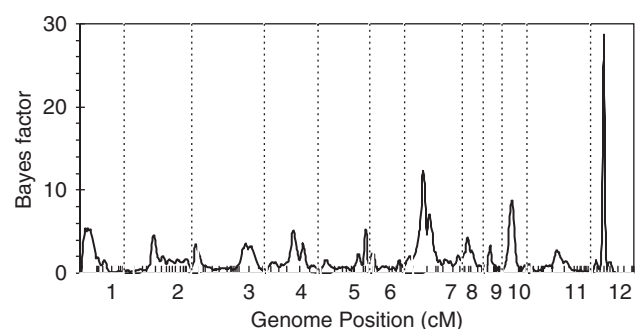and

$$\hat{\Sigma} = \begin{bmatrix} 0.5767 & 0.1231 & -0.0547 \\ 0.1231 & 0.0493 & -0.0099 \\ -0.0547 & -0.0099 & 0.0053 \end{bmatrix},$$

respectively. The estimated residual variance is $\hat{\sigma}^2 = 0.0083$.

The profiles of the BF for each locus across the genome are depicted in the top plot in Figure 2. Compared with the relative profiles in Figure 1, 12 peaks can be found, including the 9 loci detected by nonepistatic analysis. Except for the peak on chromosome 11, others show strong evidence for the presence of QTLs.

As shown in Figure 3, Bayesian epistatic analysis found that four pairs of QTLs on chromosomes 1, 2, 3 and 4 perform strong interactions, and that the QTL pair on chromosomes 3 and 10 and the one on chromosomes 4 and 8 have relatively high BF values, but the interactions are nonsignificant. Note that the fourth QTL on chromosome 3 and the eighth QTL on chromosome 8 are not found in nonepistatic analysis. Hence, we infer that the fourth and eighth QTLs are detected in epistatic analysis, mainly because of epistatic interactions.

Estimates for main-effect and for epistatic-effect QTL parameters, including QTL positions, regression effects and BFs, are shown in Tables 1 and 2, respectively. To illustrate the effects of QTLs on dynamic traits, we depict the changes in the main effects of 12 QTLs with measurement time in Figure 4. These curves are combined onto the three groups: convex (above), concave (middle) and linear (below) ones. We find that the 10th QTL and the 12th QTL on dynamic traits have strong influences on the change in direct and inverse propor-
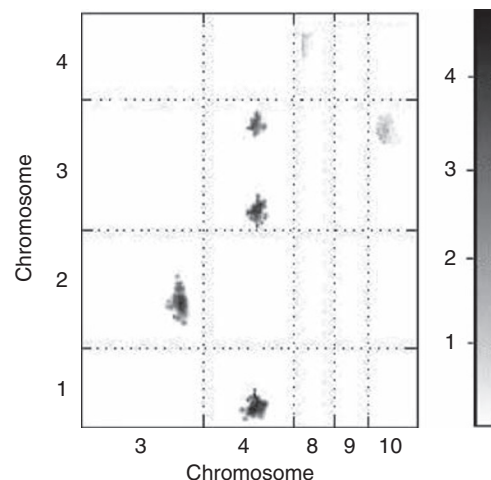

Figure 3 The two-dimensional profile of Bayes factors for epistatic effects on the selected chromosomes.

Table 1 Estimates for regression effects of main-effect QTLs detected with Bayesian epistatic mapping analysis

QTL no.	Chr-position	β_0	β_1	β_2	BF
1	1-19.6	0.171	0.054	−0.020	5.4
2	2-66.3	0.194	0.061	−0.020	4.6
3	3-13.3	−0.195	−0.009	0.027	3.5
4	3-129.4	−0.023	−0.005	0.001	3.2
5	4-67.5	−0.057	0.011	0.029	5.1
6	5-106.2	−0.010	−0.064	−0.002	5.3
7	7-44.2	−0.253	−0.054	0.024	12.3
8	8-15.1	0.156	0.027	−0.028	4.3
9	9-16.9	0.055	−0.041	0.011	3.3
10	10-23.8	0.412	0.087	−0.043	8.8
11	11-68.3	−0.110	−0.010	0.022	2.8
12	12-31.7	−0.312	−0.124	0.007	28.5

Abbreviation: QTL, quantitative trait loci.

Table 2 Estimates for regression effects of epistatic QTLs identified with Bayesian epistatic mapping analysis

QTL pair	Chr-position	δ_0	δ_1	δ_2	BF
1	1-19.6 × 4-67.5	0.184	0.026	−0.002	5.0
2	2-66.3 × 3-129.4	−0.168	−0.060	0.035	3.7
3	3-13.3 × 4-67.5	0.039	−0.022	−0.016	4.2
4	3-129.4 × 4-67.5	−0.056	−0.060	−0.026	3.5

Abbreviation: QTL, quantitative trait loci.

tion, respectively, with growth time, whereas the effects of other QTLs do not result in distinct changes.

Simulation

We simulated a dynamic trait measured at eight time points for 150 or 300 BC individuals. A genome consisting of a single large chromosome of 600 cM was simulated, which was covered by 61 evenly placed markers. The growth pattern of the dynamic trait was assumed to be controlled by the four additive QTLs and two pairs of epistatic QTLs with their positions and effects listed in Table 3. The order of the polynomial was set at 3, which generated the 'S' shape growth trajectory for phenotypes. The dynamic trait is measured at the
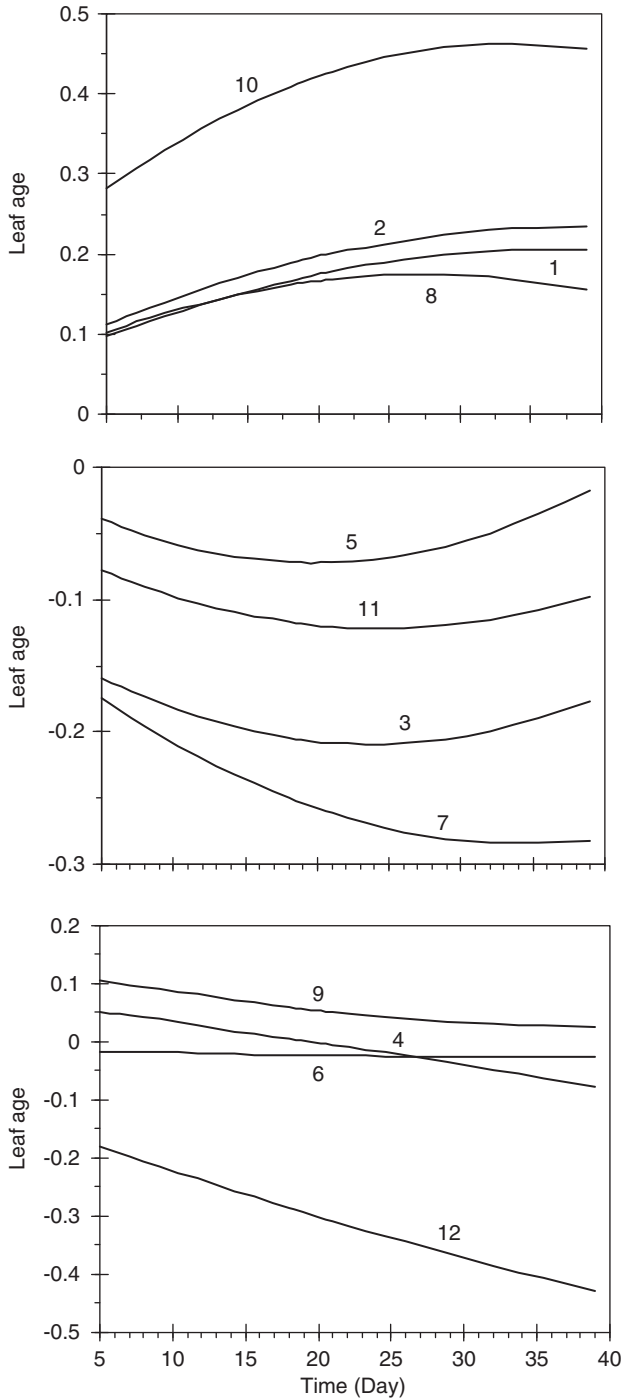
Figure 4 Changes in main effects of QTL detected with time. The above, middle and below are the convex, concave and linear groups, respectively.

Table 3 The regression effects of additive and epistatic QTLs simulated

QTL no.	QTL type	Position (cM)	β_0	β_1	β_2	β_3
1	Additive	23	0.00	1.65	2.52	1.20
2		148	2.55	1.36	−2.02	−1.27
3		332	2.94	0.00	−1.08	1.72
4		522	2.00	−1.25	0.00	−1.28
			δ_0	δ_1	δ_2	δ_3
5	Epistatic	23 × 332	1.82	−0.80	−1.20	−0.80
6		148 × 522	1.75	1.30	−1.45	1.17

Abbreviation: QTL, quantitative trait loci.

In all analyses for simulated data, we set the prior number of main-effect QTLs at 4 and the prior expected number of epistatic QTLs at 2. The upper bound of the number of QTLs was then $L = 6 + 3\sqrt{6} = 13$. The actual values for the hyperparameters used here take the same values as in real data analyses. The initial values of all variables were sampled from their prior distributions. The MCMC is run for 10 000 cycles as a burn-in period (deleted) and then for an additional 150 000 cycles after the burn-in. The chain is then thinned to reduce serial correlation by saving one observation in every 50 cycles. The posterior sample contained 3000 observations for the post-MCMC analysis. Note that here the length of the burn-in is judged by visually inspecting the plots of some posterior samples across rounds and is set to enough cycles for ensuring the MCMC convergence. The simulation experiment is replicated 40 times for evaluating the statistical power of our proposed method. The statistical power is calculated as the percentage of the number of those simulations in which significant QTL is detected.

The purpose of the simulation is to show the performance of the method proposed herein in simultaneously detecting main-effect and epistatic QTLs under different sample sizes. Therefore, we do not compare our approach with other methods for only mapping main-effect QTLs, such as the maximum likelihood approach. Table 4 shows the estimates for regression effects of the given QTLs in Table 3 and the relative statistical power of QTL detection. Apparently, Bayesian mapping of genome-wide interacting loci for dynamic traits is able to accurately estimate the regression effects of QTLs detected. Furthermore, the estimation precision of parameters and statistical power of QTL detection, as expected, improve with the increasing effect or genetic contribution proportion of QTL and increasing sample sizes. In addition, we find that the Bayesian model selection for mapping QTLs of dynamic traits is sensitive to QTLs with a relatively small genetic effect, compared with the mapping results of QTLs with the same regression effects but a lower residual variance in Yang and Xu (2007).

Discussion

By assigning a maximum number of detectable QTLs and using latent binary variables to indicate which main and epistatic effects of putative QTLs are included in or excluded from the model, Yi *et al*. (2005) first applied a

same 8 time points as in real data. The simulated population mean was $\mu = [45 \ 44 \ -1 \ -7]^T$, covariance matrix for individual-specific environmental error was

$$\Sigma = \begin{bmatrix} 1.042 & 0.171 & -0.035 & 0.100 \\ 0.171 & 0.086 & 0.041 & 0.032 \\ -0.035 & 0.041 & 0.087 & 0.052 \\ 0.100 & 0.032 & 0.052 & 0.076 \end{bmatrix}$$

and the residual variance was taken at 4.0.

Table 4 Mean estimates and s.d. (in parentheses) of QTL regression effects and statistical power of QTL detection

Sample size 150	QTL no.	Position (cM)	Estimate				Power (%)
			β_0	β_1	β_2	β_3	
	1	21.3 (3.6)	−0.06 (0.43)	1.69 (0.53)	2.48 (0.61)	1.15 (0.33)	65
	2	147.3 (2.8)	2.67 (0.23)	1.43 (0.36)	−1.96 (0.59)	−1.30 (0.53)	75
	3	334.1 (2.3)	3.04 (0.42)	0.03 (0.29)	−0.92 (0.53)	1.83 (0.51)	90
	4	524.7 (3.1)	1.93 (0.31)	−1.32 (0.51)	0.05 (0.26)	−1.19 (0.49)	85
			δ_0	δ_1	δ_2	δ_3	
	5	22.1 (3.3) × 334.7 (2.9)	1.74 (0.22)	−0.96 (0.43)	−1.16 (0.49)	−0.95 (0.63)	75
	6	147.4 (3.5) × 523.9 (2.3)	1.79 (0.38)	1.25 (0.47)	−1.37 (0.26)	1.23 (0.19)	75
300			β_0	β_1	β_2	β_3	
	1	21.9 (3.3)	0.02 (0.31)	1.66 (0.42)	2.49 (0.58)	1.18 (0.38)	75
	2	148.5 (2.7)	2.59 (0.21)	1.40 (0.33)	−2.01 (0.62)	−1.31 (0.48)	90
	3	333.7 (2.1)	2.98 (0.38)	0.02 (0.31)	−1.06 (0.46)	1.77 (0.39)	100
	4	520.5 (3.0)	2.02 (0.26)	−1.27 (0.47)	0.03 (0.26)	−1.24 (0.52)	95
			δ_0	δ_1	δ_2	δ_3	
	5	23.3 (3.2) × 332.1 (3.0)	1.87 (0.29)	−0.87 (0.33)	−1.17 (0.44)	−0.89 (0.35)	90
	6	148.6 (3.1) × 522.3 (2.3)	1.77 (0.033)	1.26 (0.29)	−1.43 (0.53)	1.19 (0.41)	85

Abbreviation: QTL, quantitative trait loci.

Bayesian model selection method to identify epistatic QTLs in experimental crosses. The approach allows MCMC sampling for QTL parameters to be carried out in the reduced model space, enhancing the computational efficiency of Bayesian mapping many epistatic QTLs. Subsequently, Yi et al. (2007a) extended a Bayesian model selection method for a single continuous trait to an ordinal trait. In this study, we adopt a multivariate version of the Bayesian model selection method to map epistatic QTL for dynamic traits. By pre-estimating indicator variables of putative QTL genotypes and exploring the posterior for indicator variables of genetic effects (Yi et al., 2007b), the Bayesian mapping method can fairly quickly identify interacting QTLs for dynamic traits in models with large numbers of genetic effects.

Generally, there are three types of epistatic interaction between QTLs: (1) where both QTLs are the main effect; (2) where both QTLs are not the main effect and (3) where only one QTL is the main effect. In mapping practice, Bayesian model selection can sensitively detect them by regulating dependence priors on genetic architecture indicators (Yi et al., 2007a, b). However, the epistatic QTLs for leaf age growth are found only between main-effect QTLs in our real data analysis.

In fact, the orders of polynomials for all effects in model (1) are unknown. We can only determine the order of polynomial for the population mean according to the shape of phenotypic trajectories of dynamic traits. In implementing our proposed method, we simple chose the Legendre polynomial functions of the same order as for population mean to fit change in QTL genetic effects and time-dependent environmental effects with time. The shape of the population mean or each effect depends on different estimates for corresponding polynomial regression coefficients. Naturally, one would ask whether the order of the Legendre polynomial for each effect is indeed the same. The choice for each submodel in model (1) will be required to answer the issue. We may first

choose the highest possible order and use it for all QTL effects and time-dependent environmental effects. For each QTL effect, we then take each regression coefficient in the nested polynomial to a different indicator variable and infer the significance of these regression coefficients by calculating the related BF value in post-MCMC analysis. For time-dependent environmental effects, however, it is difficult to infer many individual-specific regression coefficients as for QTL effects because of the large number of regression coefficients. In this case, we can adopt Bayesian model selection for random covariance matrix in mixed model (Chen and Dunson, 2003; Kinney and Dunson, 2007) to determine the order of the Legendre polynomial for time-dependent environmental effects. Once some appropriate submodels are chosen for the population mean, all QTL effects and time-dependent environmental effects by using the described procedures above and the optimal multiple interacting QTL model for dynamic traits will be established. In choosing the submodel of each QTL effect and Bayesian model selection for the random covariance matrix, the priors and posteriors for many new unknown variables need to be specified and deduced under multiple interacting QTL models for dynamic traits. These are being implemented in our research plan.

In addition, how to model residuals is also a noticeable question. Functional mapping recommended a parametric residual covariance structure by using the time series autocorrelation structure. The autoregressive model with order 1 [AR(1)] and one unknown parameter is often used in functional mapping. However, there appears to be no efficient way to sample the autoregressive coefficient in a covariance matrix within the Bayesian framework. Our investigation found that the specifying uniform distribution as a prior for autoregressive coefficient and the sampling method proposed by Gianola et al., 2003 do not work in Bayesian functional mapping. In fact, the covariance structure described by

$\psi^{\mathrm{T}}\Sigma\psi + I\sigma^2$ is more flexible than the parametric structure because we can actually choose a different degree of the polynomial to fit a covariance structure with a different degree of complexity. Moreover, we can easily sample the covariance matrix Σ from a closed form of marginal posterior distribution.

The multiple interacting QTL model for dynamic traits proposed herein can be treated as a general form of the model for analyzing the genetic architecture of continuous traits. For instance, letting $\psi = 1$ and $\xi_i = 0$ in scale, that is, only one measurement on each individual, leads to multiple interacting QTL models for single continuous quantitative traits; taking ψ to an identity matrix of m order and ξ_i to a zero vector results in a multiple interacting QTL model for multiple continuous quantitative traits; and ff ξ_i is assigned to nonzero in the two cases above. The multiple interacting QTL models for a single continuous quantitative trait and multiple continuous quantitative traits are also able to make use of repeat records on the phenotypes. Corresponding Bayesian model selection approaches can be likewise obtained by taking ψ and ξ_i to different values or matrices.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

References

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

Chen Z, Dunson DB (2003). Random effects selection in linear mixed models. *Biometrics* **59**: 762–769.

Cheverud JM, Rutledge JJ, Atchley WR (1983). Quantitative genetics of development, genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution* **37**: 895–905.

Carlin BP, Chib S (1995). Bayesian model choice via Markov chain Monte Carlo. *J Am Stat Assoc* **88**: 881–889.

Eaves LJ, Neale MC, Maes H (1996). Multivariate multipoint linkage analysis of quantitative trait loci. *Behav Genet* **26**: 519–525.

Emebiri LC, Devey ME, Matheson AC, Slee MU (1998). Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theor Appl Genet* **97**: 1053–1061.

Gelman A, Carlin JB, Stern HS, Rubin DB (1995). *Bayesian Data Analysis*. Chapman & Hall: New York.

Gianola D, Perez-Enciso M, Toro MA (2003). On marker-assisted prediction of genetic value: Beyond the ridge. *Genetics* **163**: 347–365.

Hastings WK (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**: 97–109.

Henderson CR, Kempthorne O, Searle SR, von Krosigk CM (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics* **15**: 192–218.

Huang SQ, Cui Y, Yang R (2005). Functional mapping of dynamic traits with Legendre polynomial. *Prog Nat Sci* **10**: 1183–1188.

Jiang C, Zeng ZB (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**: 1111–1127.

Kao CH, Zeng ZB (2002). Modeling epistasis of quantitative trait loci using Cockerham's model. *Genetics* **160**: 1243–1261.

Kass RE, Raftery AE (1995). Bayes factors. *J Am Stat Assoc* **90**: 773–795.

Kinney SK, Dunson DB (2007). Fixed and random effects selection in linear and logistic models. *Biometrics* **63**: 690–698.

Kirkpatrick M, Heckman N (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J Math Biol* **27**: 429–450.

Kirkpatrick M, Lofsvold D, Bulmer M (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* **124**: 979–993.

Knott SA, Haley CS (2000). Multitrait least squares for quantitative trait loci detection. *Genetics* **156**: 899–911.

Kohn R, Smith M, Chan D (2001). Nonparametric regression using linear combinations of basis functions. *Stat Comput* **11**: 313–322.

Korol AB, Ronin YI, Kirzhner VM (1995). Interval mapping of quantitative trait loci employing correlated trait complexes. *Genetics* **140**: 1137–1147.

Ma CX, Casella G, Wu RL (2002). Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. *Genetics* **61**: 1751–1762.

Macgregor S, Knott SA, White I, Visscher PM (2005). Quantitative trait locus analysis of longitudinal quantitative trait data in complex pedigrees. *Genetics* **171**: 1365–1376.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953). Equation of state calculations by fast computing machines. *J Chem Phys* **21**: 1087–1092.

Nuzhdin SV, Pasyukova EG, Dilda CL, Zeng ZB, Mackay TFC (1997). Sex-specific quantitative trait loci affecting longevity in Drosophila melanogaster. *Proc Natl Acad Sci USA* **94**: 9734–9739.

Plummer M, Best N, Cowles K, Vines K (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News* **6**: 7–10.

Raftery AE, Madigan D, Hoeting JA (1997). Bayesian model averaging for linear regression models. *J Am Stat Assoc* **92**: 179–191.

Ronin YI, Kirzhner VM, Korol AB (1995). Linkage between loci of quantitative traits and marker loci: multi-trait analysis with a single marker. *Theor Appl Genet* **90**: 776–786.

Schaeffer LR (2004). Application of random regression models in animal breeding. *Livest Prod Sci* **86**: 35–45.

Sillanpää MJ, Corander J (2002). Model choice in gene mapping: what and why. *Trends Genet* **18**: 301–307.

Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A (1997). Quantitative trait dissection analysis in Eucalyptus using RADP markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor Appl Genet* **95**: 597–608.

Wang H, Zhang YM, Li X, Masinde GL, Mohan S, Baylink DJ *et al.* (2005). Bayesian shrinkage estimation of quantitative trait loci parameters. *Genetics* **170**: 465–480.

Weng Q, Wu W, Li W, Liu H, Tang D, Zhou Y *et al.* (2000). Construction of an RFLP linkage map of rice using DNA probes from two different sources. *J Fujian Agric Univ* **29**: 129–133.

Wu R, Lin M (2006). Opinion: functional mapping—how to map and study the genetic architecture of dynamic complex traits. *Nat Rev Gen* **7**: 229–237.

Wu R, Ma CX, Lin M, Casella G (2004a). A general framework for analyzing the genetic architecture of developmental characteristics. *Genetics* **166**: 1541–1551.

Wu R, Ma CX, Lin M, Wang Z, Casella G (2004b). Functional mapping of quantitative trait loci underlying growth trajectories using a transform-both-sides logistic model. *Biometrics* **60**: 729–738.

Wu R, Ma CX, Zhu J, Casella G (2002). Mapping epigenetic quantitative trait loci (QTL) altering a developmental trajectory. *Genome* **45**: 28–33.

Wu R, Wang Z, Zhao W, Cheverud JM (2004c). A mechanistic model for genetic machinery of ontogenetic growth. *Genetics* **168**: 2383–2394.

Wu WR, Li WM, Tang DZ, Lu HR, Worland AJ (1999). Time-related mapping of quantitative trait loci underlying tiller number in rice. *Genetics* **151**: 297–303.

Yan J, Zhu J, He C, Benmoussa M, Wu P (1998a). Molecular dissection of developmental behavior of plant height in rice (Oryza sativa L.). *Genetics* **150**: 1257–1265.

Yan JQ, Zhu J, He CX, Benmoussa M, Wu P (1998b). Quantitative trait loci analysis for the developmental behavior of tiller number in rice (Oryza sativa L.). *Theor Appl Genet* **97**: 267–274.

Yang R, Gao H, Wang X, Zhang J, Zeng ZB, Wu R (2007). A semiparametric approach for composite functional mapping of dynamic quantitative traits. *Genetics* **177**: 1859–1870.

Yang R, Tian Q, Xu S (2006). Mapping quantitative trait loci for longitudinal traits in line crosses. *Genetics* **173**: 2339–2356.

Yang R, Xu S (2007). Bayesian shrinkage analysis of quantitative trait loci for dynamic traits. *Genetics* **176**: 1169–1185.

Yang RQ, Gao HJ, Sun H, Xu S (2004). Maximum likelihood analysis for mapping dynamic trait QTL in outbred population I. Methodology. *Acta Genet Sin* **31**: 1116–1122.

Yi N (2004). A unified Markov chain Monte Carlo framework for mapping multiple quantitative trait loci. *Genetics* **167**: 967–975.

Yi N, Banerjee S, Pomp D, Yandell BS (2007a). Bayesian mapping of genomewide interacting quantitative trait loci for ordinal traits. *Genetics* **176**: 1855–1864.

Yi N, Shriner D, Banerjee S, Mehta T, Pomp D, Yandell BS (2007b). An efficient Bayesian model selection approach for interacting quantitative trait loci models with many effects. *Genetics* **176**: 1865–1877.

Yi N, Yandell BS, Churchill GA, Allison DB, Eisen EJ, Pomp D (2005). Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. *Genetics* **170**: 1333–1344.

Zhou Y, Li W, Wu W, Chen Q, Mao D, Worland AJ (2001). Genetic dissection of heading time and its components in rice. *Theor Appl Genet* **102**: 1236–1242.

Appendix A

Posterior distributions for unknown parameters

The marginal posterior distribution of μ, given all other parameters, is a multivariate normal with the mean

$$(n\psi V^{-1}\psi^{\mathrm{T}})^{-1}\psi V^{-1}\sum_{i=1}^{n}(y_i - M_i + \psi^{\mathrm{T}}\mu)$$

and the covariance matrix $(n\psi V^{-1}\psi^{\mathrm{T}})^{-1}$.

The marginal posterior distribution of β_j is also a normal, of which the mean is

$$\hat{\beta}_j = \left[\left(1 + \frac{1}{c}\right)\sum_{i=1}^{n}x_{ij}^2\psi V^{-1}\psi^{\mathrm{T}}\right]^{-1}\psi V^{-1}$$

$$\times \sum_{i=1}^{n}x_{ij}(y_i - M_i + \gamma_j x_{ij}\psi^{\mathrm{T}}\beta_j)$$

and the covariance matrix is

$$\hat{\Sigma}_j = \left[\left(1 + \frac{1}{c}\right)\sum_{i=1}^{n}x_{ij}^2\psi V^{-1}\psi^{\mathrm{T}}\right]^{-1} \quad \text{for } j = 1, 2, \ldots, p$$

Likewise, the marginal posterior distribution of δ_{jk} can be expressed as a normal distribution with a mean

$$\hat{\delta}_{jk} = \left[\left(1 + \frac{1}{c}\right)\sum_{i=1}^{n}z_{ijk}^2\psi V^{-1}\psi^{\mathrm{T}}\right]^{-1}\psi V^{-1}$$

$$\times \sum_{i=1}^{n}z_{ijk}(y_i - M_i + \gamma_{jk}z_{ijk}\psi^{\mathrm{T}}\delta_{jk})$$

and the covariance matrix

$$\hat{\Sigma}_{jk} = \left[\left(1 + \frac{1}{c}\right)\sum_{i=1}^{n}z_{ijk}^2\psi V^{-1}\psi^{\mathrm{T}}\right]^{-1}$$

$$\text{for } j = 1, 2, \ldots, q; k = j+1, j+2, \ldots, q$$

The marginal posterior distribution of ξ_i subjects to normal distribution with a mean

$$\Sigma\psi V^{-1}(y_i - M_i)$$

and a covariance matrix

$$\Sigma - \Sigma\psi V^{-1}\psi^{\mathrm{T}}\Sigma$$

where the marginal posterior distribution of Σ is $\mathrm{IW}(v_a + n, \sum_{i=1}^{n}\xi_i\xi_i^{\mathrm{T}} + S_a)$.

For the residual variance σ_0^2, the corresponding marginal posterior distribution is a scaled inverse χ^2 with parameters $v_e + n$ and $(v_e + n)s_e + (\sum_{i=1}^{n}e_i^{\mathrm{T}}e_i)^{-1}$, where $e_i = y_i - M_i - \psi^{\mathrm{T}}\xi_i$.

The marginal posterior distribution of $\gamma_\bullet$ is a Bernoulli with a probability

$$p(\gamma_\bullet = 1) = \frac{wR}{(1 - w) + wR}$$

where, $w = w_m$ and $R = \sqrt{\frac{c}{c+1}}\exp(-\frac{1}{2}\hat{\beta}_j^{\mathrm{T}}\hat{\Sigma}_j^{-1}\hat{\beta}_j)$ $(j = 1, 2,\ldots,p)$ for the additive; $w = w_e$ and $R = \sqrt{\frac{c}{c+1}}\exp(-\frac{1}{2}\hat{\delta}_{jk}^{\mathrm{T}}\hat{\Sigma}_{jk}^{-1}\hat{\delta}_{jk})$ $(j = 1, 2,\ldots, q; k = j+1, j+2,\ldots, q)$ for the epistatic. The Metropolis–Hastings algorithm is also used to sample $\gamma_\bullet$ with acceptance rate $\frac{1-w}{w}\left(\frac{wR}{1-w}\right)^{1-2\gamma}$.

All aforementioned parameters have explicit forms so that samples can be directly drawn from their corresponding distributions by adopting the Gibbs sampler algorithm. The parameters without closed conditional posterior distribution forms, such as λ and X, will be sampled by using the Metropolis–Hastings algorithm. We sample QTL positions in L variable intervals whose boundaries are the positions of adjoining QTLs and restrict the minimal distance between two QTLs to be 5 cM. The Metropolis–Hastings algorithm is required to calculate an acceptance rule for accepting the proposed value over the current value. A detailed formula of the MH acceptance rule can be found for λ and X in Yang and Xu (2007).

Appendix B

Simplification of the inverse and determinant for *V*
According to the formula proved by Henderson *et al.* (1959)

$$(ZDZ^T + R)^{-1} = R^{-1} + R^{-1}Z(ZR^{-1}Z^T + D^{-1})Z^TR^{-1}$$

if we let $R = I\sigma^2$, $Z = \psi$ and $D = S$, then the inverse of V can be simplified as

$$V^{-1} = \sigma^{-2} + \sigma^{-2}\psi(\psi^T\psi\sigma^{-2} + \Sigma^{-1})\psi^T\sigma^{-2}$$

For the determinant of V,

$$
\begin{aligned}
|V| &= |\psi^T\Sigma\psi + I\sigma^2| \\
&= |I\sigma^2(\Sigma^{-1} + \psi^T\psi\sigma^{-2})\Sigma| \\
&= |I\sigma^2||\Sigma^{-1} + \psi^T\psi\sigma^{-2}||\Sigma| \\
&= \sigma^{2m}|\Sigma^{-1} + \psi^T\psi\sigma^{-2}||\Sigma|
\end{aligned}
$$

Apparently, only the inverse and determinant for $p + 1$ order matrices are required to be calculated in solving the inverse and determinant of V.