ARTICLE

Bayesian Mapping of Quantitative Trait Loci for Multiple Complex Traits with the Use of Variance Components

Jianfeng Liu, Yongjun Liu, Xiaogang Liu, and Hong-Wen Deng

Complex traits important for humans are often correlated phenotypically and genetically. Joint mapping of quantitative-trait loci (QTLs) for multiple correlated traits plays an important role in unraveling the genetic architecture of complex traits. Compared with single-trait analysis, joint mapping addresses more questions and has advantages for power of QTL detection and precision of parameter estimation. Some statistical methods have been developed to map QTLs underlying multiple traits, most of which are based on maximum-likelihood methods. We develop here a multivariate version of the Bayes methodology for joint mapping of QTLs, using the Markov chain–Monte Carlo (MCMC) algorithm. We adopt a variance-components method to model complex traits in outbred populations (e.g., humans). The method is robust, can deal with an arbitrary number of alleles with arbitrary patterns of gene actions (such as additive and dominant), and allows for multiple phenotype data of various types in the joint analysis (e.g., multiple continuous traits and mixtures of continuous traits and discrete traits). Under a Bayesian framework, parameters—including the number of QTLs—are estimated on the basis of their marginal posterior samples, which are generated through two samplers, the Gibbs sampler and the reversible-jump MCMC. In addition, we calculate the Bayes factor related to each identified QTL, to test coincident linkage versus pleiotropy. The performance of our method is evaluated in simulations with full-sib families. The results show that our proposed Bayesian joint-mapping method performs well for mapping multiple QTLs in situations of either bivariate continuous traits or mixed data types. Compared with the analysis for each trait separately, Bayesian joint mapping improves statistical power, provides stronger evidence of QTL detection, and increases precision in estimation of parameter and QTL position. We also applied the proposed method to a set of real data and detected a coincident linkage responsible for determining bone mineral density and areal bone size of wrist in humans.

In QTL mapping studies, it is a common practice to collect data observations on multiple traits. Generally, QTLs are mapped for the traits by use of single-trait analyses (i.e., analyzing the traits separately and reporting the respective results), although many of these traits are correlated phenotypically and genetically. An alternative way of QTL mapping for these correlated traits is joint analysis, by which multiple traits are analyzed simultaneously under a unified genetic model. The advantages of joint analysis over single-trait analysis are twofold. First, joint analysis may increase the statistical power and precision of QTL mapping because it combines phenotype information from all the traits and takes into account the correlation (environmental and genetic) structure of the traits. Second, genetic mechanisms of QTLs on correlated traits can be probed directly through joint analysis. For example, we can dissect a portion of genetic variation and covariation among traits by localizing QTLs and testing whether the genetic correlation is due to pleiotropy (i.e., a QTL influencing multiple correlated phenotypes) or coincident linkage (i.e., closely linked distinct QTLs for different traits).

Several methods have been developed for QTL mapping with the use of joint analysis for multiple traits.[1–14] Among these, different types of traits have been investigated, such as multiple continuous traits with normal distribution,[1,2,5–7,11,12] multiple discrete traits,[3,8] and mixtures of traits with different phenotype distributions.[4,8–10] These studies can be grouped into three broad categories according to the statistical methods adopted, each of which has its advantages and limitations.

The first method is based on the likelihood-ratio test.[1,7,8,10] A concern with this method is that increasing the number of traits may lead to a large number of unknown parameters in the model; thus, potential gains from joint analysis may be offset by the increased degrees of freedom of the test statistic.[10] In this situation, joint analysis could be less powerful than single-trait analysis.[14]

The second method adopts a semiparametric approach based on generalized estimation equations,[3] which can avoid the difficulties of full-likelihood methods and can have some advantages, such as avoidance of distribution assumptions and the ease and flexibility of model specification.[3] This method was developed under the double-cross design for experimental populations; its performance in natural populations (such as humans) remains

unclear, and its performance relative to that of the single-trait analysis is not yet known.

The third method is based on the dimension-reduction technique and involves the performance of principal-components analysis (PCA)[2] for multiple correlated traits. Analysis of the "super traits" generated via PCA requires little extra effort in comparison with single-trait analysis. However, joint analysis based on PCA may incur spurious linkage[12,15] and cause difficulties in the biological interpretation of study results.

In QTL mapping for multiple correlated traits, the Bayesian method represents another useful tool because of some of its distinct advantages over the above three methods, such as the ease of results interpretation; the inherent flexibility in handling more-complex models, various experimental populations, and various types of phenotypes; and the resultant ability to incorporate information from different sources.[16–18] The Bayesian approach has been successfully used for QTL mapping of a single trait.[16,17,19–29] However, so far, little work has been done for multiple correlated traits under a Bayesian framework.[11] In particular, no Bayesian method has been developed for mapping QTLs underlying correlated traits of a mixture of quantitative and discrete traits.

In this article, under an identity-by-descent (IBD)–based variance-components model, we developed a multivariate version of the Bayes method for QTL mapping of multiple traits under different situations, such as with multiple quantitative traits and mixed types of traits in humans. In addition, we show how to distinguish pleiotropy from coincident linkage, using the Bayes factor (BF). Our proposed method is implemented through the Markov chain–Monte Carlo (MCMC) algorithm, in which two samplers, the Gibbs sampler[30] and the reversible-jump MCMC (RJ-MCMC),[31] are employed to generate joint posterior distributions of all unknowns. We conducted two simulation experiments to evaluate the performance of our method. The results show that our joint-analysis method outperforms univariate analyses and performs well for QTL mapping of multiple correlated traits under various situations. In addition, we applied our Bayesian joint-mapping method to a set of real data on humans. The results show that the proposed method has practical significance in the human genetics field.

Methods
Bayesian Mapping for Multiple Continuous Traits

Statistical model.—Suppose there are n individuals with observations on m continuous traits in an investigated population. For ease of presentation, we adopt a full-sib design without having records for parents in each family. We assume that parents in all families come from an outbred population, where the distribution of genotypes at each locus is in Hardy-Weinberg equilibrium and the alleles of all loci under study are in linkage equilibrium. The multivariate mixed model can be written as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \sum_{i=1}^{l} \mathbf{W}\mathbf{q}_i + \mathbf{e} \ , \qquad (1)$$

where $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_k, \dots, \mathbf{y}'_n)'$ is an $nm \times 1$ vector of observations of n individuals on m traits ($\mathbf{y}_k$ is an $m \times 1$ vector of individual k on m traits), $\mathbf{X}$ is a design matrix, and β is a vector of fixed effects. In Bayesian theory, a "fixed effect" actually refers to the random effect with infinite prior variance, and its prior distribution is commonly assumed to be uniform with predefined upper and lower limits. $\mathbf{Z}$ is a design matrix relating records to individuals, $\mathbf{u}$ is a vector of additive polygenic effects, $\mathbf{W}$ is a matrix relating each individual's record to its QTL effects, $\mathbf{q}_i$ is a vector of additive QTL effects corresponding to the ith QTL, and $\mathbf{e}$ is a vector of residuals. Here, we define l, the number of QTLs, as an unknown to be estimated. In model (1), suppose that the random vectors $\mathbf{u}$, $\mathbf{q}_i$, and $\mathbf{e}$ are mutually independent, each following multivariate normal distributions as follows:

$$\mathrm{p}(\mathbf{u}\,|\,\mathbf{G},\mathbf{A}) = N(\mathbf{0},\mathbf{G}\otimes\mathbf{A}) \ , \qquad (2)$$

where $\mathbf{G}$ is the $m \times m$ (co)variance matrix of the polygenic effects across m traits, $\mathbf{A}$ is the additive genetic relationship matrix that can be obtained from the pedigree of all individuals, and $\otimes$ is the Kronecker product operator.

$$\mathrm{p}(\mathbf{q}_i\,|\,\mathbf{Q}_i,\mathbf{M},\lambda_i) = N(\mathbf{0},\mathbf{Q}_i\otimes\mathbf{\Pi}_i) \ , \qquad (3)$$

where $\mathbf{Q}_i$ is the $m \times m$ (co)variance matrix of the ith QTL effects across m traits. $\mathbf{\Pi}_i$ is the IBD matrix for the ith QTL conditional on marker data ($\mathbf{M}$) and the position (λ_i) of the ith QTL within the chromosome region considered. In this article, we use the IBD information of all markers simultaneously to infer IBD matrix $\mathbf{\Pi}_i$, on the basis of the multipoint method proposed by Xu and Gessler.[32]

$$\mathrm{p}(\mathbf{e}\,|\,\mathbf{R}) = N(\mathbf{0},\mathbf{R}\otimes\mathbf{I}) \ , \qquad (4)$$

where $\mathbf{R}$ is the $m \times m$ residual (co)variance matrix and $\mathbf{I}$ is an identity matrix.

Joint posterior distribution of parameters.—In a Bayesian framework, joint posterior distribution is the prerequisite to Bayesian inference. From model (1) and equations (2)–(4), parameters to be estimated include the number of QTLs, l; the locations of QTLs, $\lambda = (\lambda_1, \dots, \lambda_l)$; and the model effects, $\theta = (\beta,\mathbf{u},\mathbf{q},\mathbf{G},\mathbf{Q},\mathbf{R})$ (here $\mathbf{q}$ denotes $\mathbf{q}_1, \dots, \mathbf{q}_l$, and $\mathbf{Q}$ denotes $\mathbf{Q}_1, \dots, \mathbf{Q}_l$). On the basis of the Bayes theorem, the joint posterior density of all parameters (l,λ,θ), conditional on the observables $\mathbf{y}$, marker information $\mathbf{M}$, and pedigree information $\mathbf{A}$, can be written as

$$\mathrm{p}(l,\lambda,\theta\,|\,\mathbf{y},\mathbf{M},\mathbf{A}) \propto \mathrm{p}(\mathbf{y}\,|\,l,\lambda,\theta)\mathrm{p}(l,\lambda,\theta) \ . \qquad (5)$$

The first term in the right-hand side of equation (5) is the model's

likelihood conditional on all unknowns, which is assumed to be multivariate normal:

$$p(\mathbf{y}|l,\boldsymbol{\lambda},\boldsymbol{\theta}) = (2\pi)^{-\frac{n}{2}}|\mathbf{R}\otimes\mathbf{I}|^{-\frac{1}{2}}$$

$$\times \exp\left[-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u}-\sum_{i=1}^{l}\mathbf{W}\mathbf{q}_i)'\right.$$

$$\left.\times (\mathbf{R}\otimes\mathbf{I})^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u}-\sum_{i=1}^{l}\mathbf{W}\mathbf{q}_i)\right] . \quad (6)$$

The second term in equation (5) is the joint prior distribution of $(l,\boldsymbol{\lambda},\boldsymbol{\theta})$, which can be extended as

$$p(l,\boldsymbol{\lambda},\boldsymbol{\theta}) = p(l)p(\boldsymbol{\beta})p(\mathbf{G})p(\mathbf{u}|\mathbf{G})p(\mathbf{R})\prod_{i=1}^{l}$$

$$\times [p(\lambda_i)p(\mathbf{Q}_i)p(\mathbf{q}_i|\mathbf{Q}_i)] . \quad (7)$$

The prior distributions of $\mathbf{u}$ and $\mathbf{q}_i$ are given in equations (2) and (3), respectively. The fixed effects $\boldsymbol{\beta}$ are random variables with uniform prior distributions on a predefined interval in the Bayesian framework:

$$p(\boldsymbol{\beta}) \propto \text{constant} . \quad (8)$$

The number of QTL, l, can be assumed to be a truncated Poisson distribution with mean μ_l and a predefined maximum L:

$$P(l = x) = \begin{cases} \dfrac{e^{-\mu_l}\mu_l^x}{x!}, & \text{for } x = 0, 1, \dots, L \\ 0, & \text{otherwise} \end{cases} . \quad (9)$$

Since we have no information about QTL positions λ_i, a prior uniform distribution with an interval equivalent to the length of the chromosome region of interest is available for each QTL:

$$p(\lambda_i) \propto \text{constant} . \quad (10)$$

As for the prior distributions of $\mathbf{G}$, $\mathbf{Q}_i$, and $\mathbf{R}$, we assume an m-dimensional inverted Wishart distribution with p df and an $m \times m$ scale matrix $\mathbf{V}$ for all these random matrices:

$$p(\mathbf{G}) = \text{IW}_m(\mathbf{V},p) ,$$

$$p(\mathbf{Q}_i) = \text{IW}_m(\mathbf{V},p) ,$$

$$\text{and } p(\mathbf{R}) = \text{IW}_m(\mathbf{V},p) . \quad (11)$$

If we have no prior knowledge about $\mathbf{G}$, $\mathbf{Q}_i$, and $\mathbf{R}$, we can set $\mathbf{V} = 0$ and $p = -(m+1)$, such that the distributions in equation (11) become uniform for these matrices.[33]

Equations (2), (3), and (6)–(11) can together fully describe equation (5), the joint posterior distribution of all unknowns. This joint posterior is the basis for deriving the fully conditional distributions of all unknowns of interest, and further Bayesian inference for these unknowns will be based on respective fully conditional distribution via implementation of the Gibbs sampler and RJ-MCMC.

Fully conditional posterior distributions of unknowns.—Following the basic idea of deriving the fully conditional distribution of a variable,[33] we pick up those terms that include this variable from the joint posterior distribution of all parameters given by equation (5).

Specifically, the fixed-effects vector $\boldsymbol{\beta}$ has a fully conditional distribution of

$$\boldsymbol{\beta}|\mathbf{u},\mathbf{q},\mathbf{G},\mathbf{Q},\mathbf{R},\mathbf{y},l,\boldsymbol{\lambda}$$

$$\sim N[(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}^*,(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-1}] , \quad (12)$$

where $\mathbf{y}^* = \mathbf{y} - \mathbf{Z}\mathbf{u} - \sum_{i=1}^{l}\mathbf{W}\mathbf{q}_i$. The $m \times 1$ polygenic additive effects subvector, $\mathbf{u}_k(k = 1,\dots,n)$, for individual k (with the assumption that each individual has observations on all m traits) has the form of the fully conditional distribution of

$$\mathbf{u}_k|\boldsymbol{\beta},\mathbf{u}_{-k},\mathbf{q},\mathbf{G},\mathbf{Q},\mathbf{R},\mathbf{y},l,\boldsymbol{\lambda} \sim N[(\mathbf{R}^{-1}+\mathbf{G}^{-1}\mathbf{A}^{kk})^{-1}$$

$$\times (\mathbf{R}^{-1}\mathbf{y}_k^* - \sum_{j\neq k}\mathbf{G}^{-1}\mathbf{A}^{kj}\mathbf{u}_j) ,$$

$$(\mathbf{R}^{-1}+\mathbf{G}^{-1}\mathbf{A}^{kk})^{-1}] , \quad (13)$$

where $\mathbf{y}_k^* = \mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta} - \sum_{i=1}^{l}\mathbf{W}_k\mathbf{q}_i$. Here, $\mathbf{X}_k$ and $\mathbf{W}_k$ denote the respective matrices $\mathbf{X}$ and $\mathbf{W}$ with only the lines from $(k-1)m + 1$ to km included, $\mathbf{u}_{-k}$ means the vector $\mathbf{u}$ with subvector $\mathbf{u}_k$ excluded, and $\mathbf{A}^{kj}$ is the (k,j)th element of the inverse of $\mathbf{A}$. The fully conditional distribution of the $m \times 1$ effects subvector of the ith QTL, $\mathbf{q}_{i,k}(k = 1,\dots,n)$, for individual k is given as

$$\mathbf{q}_{i,k}|\boldsymbol{\beta},\mathbf{u},\mathbf{q}_{-(i,k)},\mathbf{G},\mathbf{Q},\mathbf{R},\mathbf{y},l,\boldsymbol{\lambda} \sim N[(\mathbf{R}^{-1}+\mathbf{Q}_i^{-1}\boldsymbol{\Pi}_i^{kk})^{-1}$$

$$\times (\mathbf{R}^{-1}\mathbf{y}_k^* - \sum_{j\neq k}\mathbf{Q}_i^{-1}\boldsymbol{\Pi}_i^{kj}\mathbf{q}_{i,j}),$$

$$(\mathbf{R}^{-1}+\mathbf{Q}_i^{-1}\boldsymbol{\Pi}_i^{kk})^{-1}] , \quad (14)$$

where $\mathbf{q}_{-(i,k)}$ denotes vector $\mathbf{q}$ with the subvector $\mathbf{q}_{i,k}$ excluded, and $\mathbf{y}_k^* = \mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta} - \mathbf{Z}_k\mathbf{u} - \sum_{i'\neq i}\mathbf{W}_k\mathbf{q}_{i'}$. The fully conditional distributions of the (co)variance matrices $\mathbf{Q}_i(i = 1,\dots,l)$, $\mathbf{G}$, and $\mathbf{R}$ are m-dimensional inverted Wishart distributions, each having $n - m - 1$ df:

$$\mathbf{Q}_i|\mathbf{q}_i,\boldsymbol{\Pi}_i \sim \text{IW}_m(\mathbf{V}_{\mathbf{Q}_i},n-m-1) ,$$

$$\mathbf{G}|\mathbf{u},\mathbf{A} \sim \text{IW}_m(\mathbf{V}_{\mathbf{G}},n-m-1) ,$$

$$\text{and } \mathbf{R}|\mathbf{e} \sim \text{IW}_m(\mathbf{V}_{\mathbf{R}},n-m-1) , \quad (15)$$

where $m \times m$ scale matrices $\mathbf{V}_{\mathbf{Q}_i}$, $\mathbf{V}_{\mathbf{G}}$, and $\mathbf{V}_{\mathbf{R}}$ act as the hyperparameters of inverted Wishart distributions corresponding to the three random matrices. Here, the (r,s)th element of matrix $\mathbf{V}_{\mathbf{Q}_i}$ is $\mathbf{V}_{\mathbf{Q}_i}(r,s) = [\mathbf{q}_{i(r)}'\boldsymbol{\Pi}_i^{-1}\mathbf{q}_{i(s)}]^{-1}$, with $\mathbf{q}_{i(r)}$ indicating the $n \times 1$ vector of the ith QTL effects for trait $r(r = 1,\dots,m)$. Similarly, the (r,s)th element of matrix $\mathbf{V}_{\mathbf{G}}$ is $\mathbf{V}_{\mathbf{G}}(r,s) = [\mathbf{u}_{(r)}'\mathbf{A}^{-1}\mathbf{u}_{(s)}]^{-1}$, and the (r,s)th element of matrix $\mathbf{V}_{\mathbf{R}}$ is $\mathbf{V}_{\mathbf{R}}(r,s) = [\mathbf{e}_{(r)}'\mathbf{e}_{(s)}]^{-1}$, with $\mathbf{e}_{(r)}$ indicating the $n \times 1$ vector of environmental effects for trait r, where $\mathbf{e}_{(r)} = \mathbf{y}_{(r)} - \boldsymbol{\beta}_{(r)} - \mathbf{u}_{(r)} - \sum_{i=1}^{l}\mathbf{q}_{i(r)}$.

Unlike the above parameters, there is no explicit expression for the fully conditional posterior distributions of the parameters l and λ. Hence, we cannot obtain the kernel of a standard density—for example, normal distribution—for either of them. Accordingly, we use RJ-MCMC, an extension of the Metropolis-Hastings sampler,[34,35] to jointly draw the samples of these two parameters with no need for the explicit expressions of their fully conditional distributions (see detailed description in the following section).

Bayesian inference of unknowns based on MCMC process.—We simulated an MCMC process for sampling the joint posterior distribution of all unknowns given in equation (5), and these joint posterior samples were used to perform Bayesian mapping and for parameter inference. The basic idea of the MCMC process is to mimic a random walk in the space of all unknowns in accordance with their respective fully conditional distributions. The simulated random walk finally converges to a stationary distribution,[36] which can be regarded as the joint posterior distribution of the unknowns.

When MCMC is performed, those unknowns except l and λ, each having the standard form of a fully conditional distribution, will be sampled through the Gibbs sampler, whereas changing the number of QTLs, l, requires a corresponding change in the dimension of the model and thus needs a reversible-jump step. This has been achieved via RJ-MCMC in suites of studies on single-trait QTL mapping.[16,17,28,29] In this article, we directly extend RJ-MCMC to sample the number of QTLs, l, and their positions, λ, in the case of multiple-trait mapping. Specifically, the MCMC process for generating the joint posterior samples of all unknowns can be briefly summarized as follows:

1. Initiate (l,λ,θ) in the space of all parameters with proper starting values in terms of their prior distributions given in equations (9) (l); (10) (λ); and (2)–(4), (8), (10), and (11) (θ).
2. Calculate IBD matrices for the corresponding QTLs in terms of their positions, using the multipoint method.[32] Sample all unknowns included in θ in turn, in accordance with their fully posterior distributions given in equations (12)–(15).
3. Jointly sample l and λ, using the updating scheme similar to those in single-trait analyses,[16,17,28,29] in accordance with three types of moves with proposal probabilities $p_m = 1/3$, $p_d = 1/3$, and $p_a = 1/3$ to (1) modify λ and keep l unchanged, (2) randomly remove one of the existing QTLs from the model if $l > 0$, and (3) add a new QTL to the model if $l < L$, respectively.
   a. When the first type of move is chosen, we update positions for existing QTLs sequentially. Specifically, for the ith QTL, a proposal position λ_i^* is sampled from a uniform distribution with symmetric interval $\lambda_i - d$ to $\lambda_i + d$ around the previous position λ_i, where d is the predefined tuning parameter. The new vector of ith QTL effects, $\mathbf{q}_i^*$, is also sampled from equation (14) in accordance with the IBD matrix $\mathbf{\Pi}_i^*$ corresponding to a new position λ_i^*. The proposal position λ_i^* is accepted with probability

$$\alpha(\lambda_i,\lambda_i^*) = \min\left\{1, \frac{\exp\left[-\frac{1}{2}\mathbf{y}_m^{*\prime}(\mathbf{R}\otimes\mathbf{I})^{-1}\mathbf{y}_m^*\right]}{\exp\left[-\frac{1}{2}\mathbf{y}^{*\prime}(\mathbf{R}\otimes\mathbf{I})^{-1}\mathbf{y}^*\right]}\right\},$$

   where $\mathbf{y}_m^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \mathbf{Wq}_i^* - \sum_{i'\neq i}\mathbf{Wq}_{i'}$ and $\mathbf{y}^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \sum_{i'=1}^{l}\mathbf{Wq}_{i'}$. If the new position is accepted, the corresponding effect vector of the ith QTL, $\mathbf{q}_i^*$, is also updated. Otherwise, both the position and effects of ith QTL remain unchanged. This sampling process will be repeated at the position of each existing QTLs.
   b. When p_d happens, each existing QTL has an equal probability of being removed from the current model. Suppose the ith QTL is proposed to be deleted; then, the multi-

variate version of acceptance probability is similar to single-trait mapping[37,38]:

$$\alpha(l,l-1)$$
$$= \min\left\{1, \frac{\exp\left[-\frac{1}{2}\mathbf{y}_d^*(\mathbf{R}\otimes\mathbf{I})^{-1}\mathbf{y}_d^*\right]}{\exp\left[-\frac{1}{2}\mathbf{y}^{*\prime}(\mathbf{R}\otimes\mathbf{I})^{-1}\mathbf{y}^*\right]} \times \frac{l}{\mu_l} \times \frac{p_a}{p_d}\right\},$$

   where $\mathbf{y}_d^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \sum_{i'\neq i}\mathbf{Wq}_{i'}$ and $\mathbf{y}^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \sum_{i'=1}^{l}\mathbf{Wq}_{i'}$. If the deletion of the ith QTL is accepted, the number of QTLs becomes $l - 1$; otherwise, the previous value l remains unchanged.
   c. If the $(l + 1)$th QTL is proposed to be added into the model, we will randomly assign starting values for corresponding parameters of the new QTL (λ_{l+1}, $\mathbf{q}_{l+1}$, and $\mathbf{Q}_{l+1}$), which is similar to step 1. When p_a happens, more QTL parameters with new starting values will be introduced to the model compared with single-trait mapping. To enhance the speed of convergence and to improve mixing of the Markov chain, we use the Gibbs sampler to perform more than one update of $\mathbf{q}_{l+1}$ and $\mathbf{Q}_{l+1}$ on the basis of alternating equations (14) and (15), before calculating the acceptance probability of adding a new QTL. In our studies, we empirically determine the extra 10 updates of parameters for the proposed QTL. Hence, the acceptance probability for the new QTL with updated parameters instead of their original starting values is given as

$$\alpha(l,l+1)$$
$$= \min\left\{1, \frac{\exp\left[-\frac{1}{2}\mathbf{y}_a^{*\prime}(\mathrm{R}\otimes\mathbf{I})^{-1}\mathbf{y}_a^*\right]}{\exp\left[-\frac{1}{2}\mathbf{y}^{*\prime}(\mathbf{R}\otimes\mathbf{I})^{-1}\mathbf{y}^*\right]} \times \frac{\mu_l}{l+1} \times \frac{p_d}{p_a}\right\},$$

   where $\mathbf{y}_a^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \sum_{i=1}^{l+1}\mathbf{Wq}_i$ and $\mathbf{y}^* = \mathbf{y} - \mathbf{X}\beta - \mathbf{Zu} - \sum_{i=1}^{l}\mathbf{Wq}_i$. If the proposed QTL is accepted, the responding QTL parameters are also accepted, and the number of QTLs becomes $l + 1$. Otherwise, the QTL number remains l.
4. The above steps constitute one iteration. The Markov chain will be constructed by repeating steps 2–3.

When large enough joint posterior samples of all unknowns are generated via the above MCMC process, QTL mapping and parameter estimation can be performed according to the marginal posterior density of all unknowns based on corresponding samples. Here, we adopt a histogram kernel to construct a QTL intensity function, which is used to determine the number and the locations of QTLs. Other parameters are also estimated from those samples in which QTL locations fall into the regions with sufficiently high estimated QTL intensities. This method has been commonly used in single-trait mapping.[16,17,19,28,29]

Bayesian Mapping for Mixed Types of Traits

In practical research, investigators may collect data of different types for a sample set. For example, both binary traits (e.g., disease status) and continuous traits (e.g., quantitative measures that are risk factors for the disease under study) are available in a study. In this study, we describe—for the first time, to our knowledge—a Bayesian method for joint mapping of loci that affect the two different types of traits. We then derive a general version of the

joint-mapping method for mixed types of multiple continuous traits and multiple ordered categorical traits.

Bivariate analysis with one binary trait and one continuous trait.—For a record in a binary fashion, an underlying normal variable, ξ_k, is often used to determine the value of w_k for individual k ($k = 1, \ldots, n$)—that is, $w_k = 1$ if $\xi_k > 0$, and $w_k = 0$ if $\xi_k \leq 0$. Because we set the threshold $t = 0$, a constraint—that residual error variance $\sigma_{e,\xi}^2 = 1$—must be superimposed to avoid overparameterization. From this threshold model, we can jointly analyze the continuous trait $\mathbf{Y}$ and the liability ξ on the basis of exactly the same methodology as the aforementioned Bayesian mapping for multiple continuous traits if only ξ is available. Hence, the key issue here is a need to generate the fully conditional distribution of the liability ξ. Following Sorensen's method,[33] we can derive the fully conditional distribution of ξ_k for individual k:

$$p(\xi_k | l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{w}, \xi_{-k}, \mathbf{y}) \propto p(\xi_k, y_k | l, \boldsymbol{\lambda}, \boldsymbol{\theta}) p(w_k | \xi_k)$$

$$\propto p(\xi_k | y_k, l, \boldsymbol{\lambda}, \boldsymbol{\theta})[\mathrm{I}(\xi_k > 0)\mathrm{I}(w_k = 1)$$

$$+ \mathrm{I}(\xi_k \leq 0)\mathrm{I}(w_k = 0)] , \qquad (16)$$

where parameters l, $\boldsymbol{\lambda}$, and $\boldsymbol{\theta}$ have the same meanings as mentioned above; the $n \times 1$ vectors $\mathbf{w} = (w_1, \ldots, w_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ contain phenotypic records for n individuals; and ξ_{-k} is the vector of liability ξ with ξ_k excluded. Note that $p(\xi_k | l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{w}, \xi_{-k}, \mathbf{y})$ is a truncated (> 0 if $w_k = 1$ and < 0 if $w_k = 0$) normal distribution from equation (16). The mean of the conditional normal distribution of ξ_k is

$$\mathrm{E}(\xi_k | l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{w}, \xi_{-k}, \mathbf{y}) = \beta_{k,\xi} + u_{k,\xi} + \sum_{i=1}^{l} q_{i,(k,\xi)}$$

$$+ \frac{-x_2}{x_1^{-1} + x_2^2}[y_k - \beta_{k,y} - u_{k,y} - \sum_{i=1}^{l} q_{i,(k,y)}] , \qquad (17)$$

and the variance is

$$\mathrm{Var}(\xi_k | l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{w}, \xi_{-k}, \mathbf{y}) = \frac{x_1^{-1}}{x_1^{-1} + x_2^2} . \qquad (18)$$

Here, the new parameters x_1 and x_2 are introduced for the drawing of samples of ξ, which are also used to update the (co)variance matrix of residual effects $\mathbf{R}$ between two traits.[33] A brief description of drawing samples of these two parameters is given below.

As mentioned earlier, we predefine the constraint $\sigma_{e,\xi}^2 = 1$ because of the feature of binary data, rather than assigning an original inverse Wishart prior to $[\sigma_{e,y}^2, \sigma_{e,y\xi}, \sigma_{e,\xi}^2]$, as shown in equation (11). Accordingly, the 2×2 (co)variance matrix $\mathbf{R}$ between two traits has the form

$$\mathbf{R} = \begin{pmatrix} \sigma_{e,y}^2 & \sigma_{e,y\xi} \\ \sigma_{e,y\xi} & \sigma_{e,\xi}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{e,y}^2 & \sigma_{e,y\xi} \\ \sigma_{e,y\xi} & 1 \end{pmatrix} ,$$

which results in a conditional inverse Wishart prior to $\sigma_{e,y}^2, \sigma_{e,y\xi} | \sigma_{e,\xi}^2 = 1$. Subsequently, the fully conditional distribution of $\sigma_{e,y}^2, \sigma_{e,y\xi} | \sigma_{e,\xi}^2 = 1, l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{y}, \xi$ follows a conditional inverse Wishart distribution. The parameters x_1 and x_2 can be used to indirectly sample $\sigma_{e,y}^2, \sigma_{e,y\xi}$ from the conditional inverse Wishart distribution.

For ease of presentation, we rewrite the scale matrix $\mathbf{V_R}$ in the inverse Wishart distribution given by equation (15):

$$\mathbf{V_R} = \begin{pmatrix} \mathbf{e}'\mathbf{e}_y & \mathbf{e}'\mathbf{e}_\xi \\ \mathbf{e}'\mathbf{e}_\xi & \mathbf{e}'\mathbf{e}_\xi \end{pmatrix}^{-1} \Rightarrow \mathbf{V_R}^{-1} = \begin{pmatrix} \mathbf{e}'\mathbf{e}_y & \mathbf{e}'\mathbf{e}_\xi \\ \mathbf{e}'\mathbf{e}_\xi & \mathbf{e}'\mathbf{e}_\xi \end{pmatrix}$$

$$= \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12} & \Psi_{22} \end{pmatrix} . \qquad (19)$$

First, we generate x_1 and x_2 by drawing successively from a one-dimensional Wishart distribution and a normal distribution, respectively:

$$(x_1 | \Psi_{11}) \sim \mathrm{W}_1(\Psi_{11}, n - 3)$$

and

$$(x_2 | x_1, \mathbf{V_R}) \sim N\left(\frac{\Psi_{12}}{\Psi_{11}}, \ \frac{|\mathbf{V_R}^{-1}|}{x_1 \Psi_{11}} \right) .$$

Second, $(x_1^{-1} + x_2^2, -x_2)$ acts as a realized value from the fully conditional distribution $\sigma_{e,y}^2, \sigma_{e,y\xi} | \sigma_{e,\xi}^2 = 1, l, \boldsymbol{\lambda}, \boldsymbol{\theta}, \mathbf{y}, \xi$. x_1 and x_2 jointly contribute to the updating of liability ξ through equations (17) and (18).

In brief, we need to make the following minor modification to the original MCMC process, to accommodate Bayesian mapping for a mixed type of bivariate analysis:

1. In step 1, we assign starting values to the liability ξ on the basis of the truncated normal prior.
2. In step 2, we generate x_1 and x_2 after updating the scale matrix $\mathbf{V_R}$ in equation (19), and then the liability ξ and the (co)variance matrix $\mathbf{R}$ are updated by x_1 and x_2. The original fully conditional distribution $\mathbf{R} | \mathbf{e} \sim \mathrm{IW}_m(\mathbf{V_R}, n - m - 1)$ for sampling $\mathbf{R}$ will no longer play any role.

Except for the above two steps, no more modifications are required to construct the chain in joint analyses of mixed types of traits.

Additionally, we have extended the above methodology to a general situation in which multiple continuous traits and multiple ordered categorical traits are considered. A detailed description for this extension is provided in appendix A.

Testing Pleiotropy versus Coincident Linkage

In analyses of multiple traits, it is important to discriminate between pleiotropy and coincident linkage. In our studies, we developed a novel method that uses the BF to test for pleiotropy versus coincident linkage. For simplicity, we detail our method for bivariate analyses here.

For model (1) mentioned earlier in this section, we assume a pleiotropic QTL model—that is, l QTLs, each affecting all traits. Suppose we have detected a QTL within the chromosomal interval k_a to k_b with sufficient posterior density. To test pleiotropy versus coincident linkage, we propose a null model involving two closely linked QTLs within the interval k_a to k_b. It is reasonable that the covariance of QTL effects on the two traits can be constrained to 0.[6,10] Thus, testing pleiotropy versus coincident linkage is the comparison of the original pleiotropic model (model$_{\text{plei.}}$)

with the coincident linkage model (model$_{link.}$), with use of the BF in a Bayesian framework.

$$\text{BF} = \frac{p(\mathbf{y} | \text{model}_{\text{plei.}})}{p(\mathbf{y} | \text{model}_{\text{link.}})} \ .$$

The BF can be calculated via harmonic mean estimators[39] of the marginal probabilities, $p(\mathbf{y} | \text{model}_{\text{plei.}})$ and $p(\mathbf{y} | \text{model}_{\text{link.}})$, given the respective model.

$$p(\mathbf{y} | \text{model}_\tau) = \frac{N}{\sum\limits_{\theta=1}^{N} \left\{ \frac{1}{p[\mathbf{y} | \mathbf{Q}_\tau^{(\theta)}, \text{else}^{(\theta)}]} \right\}} \ ,$$

where τ denotes either the pleiotropic model or the coincident linkage model fitted, and N is the number of total iterations in the MCMC process. $\mathbf{Q}_\tau^{(\theta)}$ and $\text{else}^{(\theta)}$ are samples of all unknowns drawn from the θth iteration. Here, $\mathbf{Q}_\tau$ is the 2×2 (co)variance matrix of the QTL, which has been located in the interval k_a to k_b, and "else" denotes samples of all unknowns except $\mathbf{Q}_\tau$. Calculation of $p[\mathbf{y} | \mathbf{Q}_\tau^{(\theta)}, \text{else}^{(\theta)}]$ is straightforward according to equation (6).

Since calculation of the BF is based on the outcomes of original joint-mapping analyses, the estimates of QTL number and the CIs of respective QTLs will be treated as constant when pleiotropy versus coincident linkage is being tested via the BF, and we only need to update the location of each QTL and its respective parameters within its interval rather than add or delete a QTL via the reversible-jump step in the MCMC process. Furthermore, because the covariance of QTLs between two traits is constrained to 0 in the coincident linkage model, a simple conditional inverse Wishart distribution[33] is used to draw the samples of $\mathbf{Q}_\tau$ under the null model, instead of using the original fully conditional distribution shown in equation (15).

On the basis of the output of the BF, we adopt the commonly used criteria[22,40] to judge which model—that is, pleiotropic or coincident linkage—is more favorable. Here, BF > 10 and 10 > BF > 3 suggest a pleiotropic model with strong and moderate evidence, respectively. BF < 0.1 and 0.1 < BF < 0.3 suggest a coincident linkage model with strong and moderate evidence, respectively. And, 0.3 < BF < 3 does not favor either model—that is, we cannot discriminate pleiotropic effects from coincidently linked QTLs within interval k_a to k_b.

Results

Simulation Studies

Design of the simulation experiments.—To evaluate the proposed method, we conducted simulations for joint mapping of multiple traits in different scenarios. For conciseness, only two traits were considered in each of the two situations (two quantitative traits involved in the first simulation and a mixture of a quantitative trait and a binary trait in the second simulation). For ease of comparison, the same set of simulated data is used in both simulations, including the population, marker, QTL information, and phenotypic data.

We simulated 1,000 full-sib families, each with 3 siblings (a total of 3,000 individuals). We assumed a chromosome region of 100 cM, with 11 evenly spaced codom-

inant markers and an intermarker distance of 10 cM. Four alleles were randomly simulated with equal probability at each marker locus. Within the region, two sets of QTLs were simulated. The first set has two QTLs both residing at 25 cM, one with variance $\sigma_{q1,1}^2 = 0.2$ for trait 1 and the other with $\sigma_{q1,2}^2 = 0.3$ for trait 2; the covariance between two traits that is due to these two QTLs is $\sigma_{q1,12} = 0$, which implies two coincidently linked QTLs. The second set has one QTL residing at 75 cM with parameters similarly defined as $\sigma_{q2,1}^2 = 0.2$ for trait 1 and $\sigma_{q2,2}^2 = 0.3 + \varepsilon$ for trait 2 and rendering covariance $\sigma_{q2,12} = (\sigma_{q2,1}^2 \times \sigma_{q2,2}^2)^{0.5}$ for the two traits, which implies a pleiotropic QTL. Here, we arbitrarily use a small positive constant, $\varepsilon = 0.01$, to make the (co)variance matrix of the second QTL have the inverse. We set the polygenic (co)variances $\sigma_{a,1}^2 = 0.6$, $\sigma_{a,2}^2 = 0.4$, and $\sigma_{a,12} = 0.2$ and residual (co)variances $\sigma_{e,1}^2 = 1.0$, $\sigma_{e,2}^2 = 1.0$, and $\sigma_{e,12} = 0$ for the two traits.

In the first simulation experiment, normal distributed data of the two traits $\mathbf{y}_1$ and $\mathbf{y}_2$ for individual $k(k = 1, \dots, 3,000)$ are generated using the infinitesimal model[41]:

$$\begin{pmatrix} y_{1,k} \\ y_{2,k} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} q_{1,1,kp} \\ q_{2,1kp} \end{pmatrix} + \begin{pmatrix} q_{1,1,kp} \\ q_{2,1,km} \end{pmatrix} + \begin{pmatrix} q_{1,2,kp} \\ q_{2,2,kp} \end{pmatrix} + \begin{pmatrix} q_{1,2,km} \\ q_{2,2,km} \end{pmatrix}$$
$$+ 0.5 \begin{pmatrix} a_{1,kp} \\ a_{2,kp} \end{pmatrix} + 0.5 \begin{pmatrix} a_{1,km} \\ a_{2,km} \end{pmatrix} + \begin{pmatrix} a_{1,kw} \\ a_{2,km} \end{pmatrix} + \begin{pmatrix} e_{1,k} \\ e_{2,k} \end{pmatrix} \ ,$$

where we take the overall mean vector $(\mu_1, \mu_2)' = (0,0)'$, and paternal- and maternal-allele effect vectors of the first QTL for individuals $k, (q_{1,1,kp}, q_{2,1,kp})$ and $k, (q_{1,1,km}, q_{2,1,km})$ are sampled from the same two-dimensional normal distribution with a 2×1 mean vector $(0, 0)'$ and a 2×2 variance matrix $0.5\mathbf{V}_{q1}$:

$$N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5\mathbf{V}_{q1} \right] = N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5 \begin{pmatrix} \sigma_{q1,1}^2 & \sigma_{q1,12} \\ \sigma_{q1,12} & \sigma_{q1,2}^2 \end{pmatrix} \right]$$
$$= N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5 \begin{pmatrix} 0.2 & 0 \\ 0 & 0.3 \end{pmatrix} \right] \ .$$

Similarly, for the second QTL, vectors $(q_{1,2,kp}, q_{2,2,kp})$ and $(q_{1,2,km}, q_{2,2,km})$ are sampled from the same two-dimensional normal distribution with 2×1 mean vector $(0, 0)'$ and 2×2 variance matrix $0.5\mathbf{V}_{q2}$:

$$N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5\mathbf{V}_{q2} \right] = N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5 \begin{pmatrix} \sigma_{q2,1}^2 & \sigma_{q2,12} \\ \sigma_{q2,12} & \sigma_{q2,2}^2 \end{pmatrix} \right]$$
$$= N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0.5 \begin{pmatrix} 0.2 & \sqrt{0.2 \times 0.3} \\ \sqrt{0.2 \times 0.3} & 0.3 + 0.01 \end{pmatrix} \right] \ .$$

The vectors $(a_{1,kp}, a_{2,kp})'$ and $(a_{1,km}, a_{2,km})'$ are paternal and maternal polygene additive effects vectors, respectively, which are sampled from

$$N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Va_1 & Va_{1,2} \\ Va_{1,2} & Va_2 \end{pmatrix} \right] = N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{pmatrix} \right] \ .$$

The vector of within-family deviation in polygene effects $(a_{1,kw}, a_{2,kw})'$ is sampled from

$$N\left[\begin{pmatrix}0\\0\end{pmatrix}, 0.5\begin{pmatrix}\sigma_{a,1}^2 & \sigma_{a,12}\\ \sigma_{a,12} & \sigma_{a,2}^2\end{pmatrix}\right] = N\left[\begin{pmatrix}0\\0\end{pmatrix}, 0.5\begin{pmatrix}0.6 & 0.2\\ 0.2 & 0.4\end{pmatrix}\right].$$

The residual effect vector, $(e_{1,k}, e_{2,k})'$, is sampled from a two-dimensional standard normal distribution. On the basis of these two simulated observation vectors, we obtained the realized phenotypic (co)variance matrix,

$$\text{Var}\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix} = \begin{pmatrix}1.9823 & 0.4120\\ 0.4120 & 2.0618\end{pmatrix},$$

which is highly consistent with the sum of all (co)variance matrices of all effects involved in the simulation.

In the second simulation experiment, we kept the original phenotypes of $\mathbf{y}_1$ the same but truncated phenotypic values of $\mathbf{y}_2$ to generate a new binary trait, $\mathbf{w}_2$, using a threshold of 0. Because we set the overall mean to be 0 for both traits, this led to 50% population incidence for $\mathbf{w}_2$.

To compare the efficiency of the joint analyses with single-trait analyses, separate analysis on each trait was also performed in the two simulation experiments. Our Bayesian mapping method for joint analyses can be directly used for single-trait analyses by taking the trait number $m = 1$. In both experiments, we assume that phenotype records are not available for all the parents.

In each MCMC analysis, we ran a single long chain. We set the burn-in iterations to 1,000, and then the next 5×10^5 cycles of simulations were performed to generate samples of all unknowns. We stored one iteration in every 50 MCMC cycles to reduce serial correlation between the samples, so that the total number of samples kept in each analysis was 10^4. For each MCMC process, the initial values for all unknowns are randomly generated within the parameter space in accordance with their prior distributions, rather than with use of predetermined specific values. The prior for the overall mean was uniform over the range -2.0 to 2.0. The priors for all variance parameters were generated from the uniform range $0–2.0$, and, for covariance parameters (only for joint analyses), the priors were generated from the range -2.0 to 2.0; the starting value for each (co)variance matrix must assure the existence of its inverse. We set the maximum number of QTLs as $L = 4$ within the chromosomal region of 100 cM in length. The initial value of l was randomly sampled from its prior distribution given by equation (9). In addition, when we updated the location of the existing ith QTL, the tuning parameter of the proposal distribution, d_i, was set to 2 cM.

Results of the simulation experiments.—The approximate Bayesian posterior distributions of QTL intensities obtained from both simulation experiments are presented in figure 1. The QTL intensity graphs corresponding to different analysis strategies, including joint mapping (de-

noted as J-12) and separate mapping for trait 1 (denoted as S-1) and trait 2 (denoted as S-2), are largely concentrated around the true locations of the two simulated QTLs. Specifically, when two continuous traits were considered in the analysis, two QTL intensity peaks (as shown in fig. 1a) were shaped on chromosomal positions of 24–25 cM and 75–76 cM for J-12, 24–25 cM and 76–77 cM for S-1, and 25–26 cM and 73–74 cM for S-2, whereas, when one continuous trait and one binary trait were involved in the analysis, two peaks (as shown in fig. 1b) were at 26–27 cM and 74–75 cM for J-12 and at 25–26 cM and 72–73 cM for S-2. (The graph for S-1 here is the same as that shown in fig. 1a because exactly the same data are adopted in both experiments.) Overall, all six graphs have two peaks that are quite close to the true positions of two QTLs. These results highly favor the two-QTL model, consistent with the cases in simulations. When comparing the shapes for J-12 with those for S-1 and S-2 in each experiment, all the highest QTL intensity peaks occurred in joint-mapping analyses, indicating that joint-mapping analyses provided stronger evidence and more information for QTL detection than did single-trait analyses. In addition, by comparing the shapes for J-12 in figure 1a and 1b, we found that some information was lost when binary data were involved in joint-mapping analyses. This is reflected by the result that the two QTL peaks for J-12 in figure 1b (for the mixture of one continuous trait and one binary trait) are less intense than those in figure 1a (for the two continuous traits).

The approximate posterior distributions of the QTL number generated from each analysis are presented in table 1. The posterior expectations, treated as estimates of the QTL number, are largely equal among different analyses and are consistent with the simulated number of QTLs. Furthermore, the posterior modes of the QTL number obtained from all analyses are the same as the true number of QTLs in both experiments. Within our expectation, the posterior variance of QTL number for multiple-trait analyses is smaller than those for single-trait analyses in both simulation experiments, and the variance of QTL number appeared smaller in the joint analysis for continuous traits than for mixed types of traits.

Besides the estimate of QTL number given above, we also obtained estimates of all other unknowns involved in model (1) from respective posterior samples. The results from different simulation experiments, given in tables 2 and 3, are detailed below.

For all analyses in the two experiments, we chose a suitable length of a chromosomal region around each QTL intensity peak as the location interval of the detected QTL, and, thus, the summed QTL intensity within the chosen region was sufficiently high. The posterior means of the QTL locations and the estimates of QTL (co)variances among traits were also obtained from the respective posterior samples, in which QTL locations fell into the chosen chromosomal region for each detected QTL. From the results shown in tables 2 and 3, we found that the QTL

interval showed a smaller length but a higher summed QTL intensity for joint mapping compared with separate mapping in both simulations. This demonstrates that joint mapping can provide stronger evidence and produce better precision for QTL detection. From the posterior means and posterior mean squared errors of QTL locations, the estimated locations were very close to the corresponding simulated true ones. However, separate analyses usually generated larger posterior mean squared errors than did joint analyses.

Similarly, QTL (co)variance estimates were largely close to the true values simulated, and the larger posterior mean squared errors of QTL variances always occurred in single-trait analyses compared with joint analyses. Note that estimates of QTL covariance among traits are available only in joint mapping for both simulations. For the first simulated set of QTLs under the coincident linkage model, the estimates of respective QTL covariance are close to 0 for both experiments. For the second simulated set of

Table 1. Inferred Posterior Distribution and Posterior Mean of the Number of QTLs in Joint Mapping and Single-Trait Mapping from the Two Simulation Experiments

Experiment and Method	Frequency						
	No. of QTLs					Mean	Variance
	0	1	2	3	4		
1[a]:							
J-12	.0002	.0054	.9928	.0012	.0004	1.9962	.0090
S-1	.0039	.0530	.8126	.1203	.0102	2.0799	.2233
S-2	.0009	.0616	.7989	.1320	.0066	2.0818	.2169
2[b]:							
J-12	.0003	.0021	.9053	.0609	.0314	2.1210	.1752
S-1	.0039	.0530	.8126	.1203	.0102	2.0799	.2233
S-2	.0025	.0234	.8325	.1187	.0229	2.1361	.2252

[a] In the first experiment, two continuous traits, y_1 and y_2, are analyzed by joint mapping (J-12) and by separate mapping for y_1 (S-1) and y_2 (S-2).

[b] In the second experiment, the continuous trait y_1 and the binary trait w_2 are analyzed by joint mapping (J-12) and by separate mapping for y_1 (S-1) and w_2 (S-2).
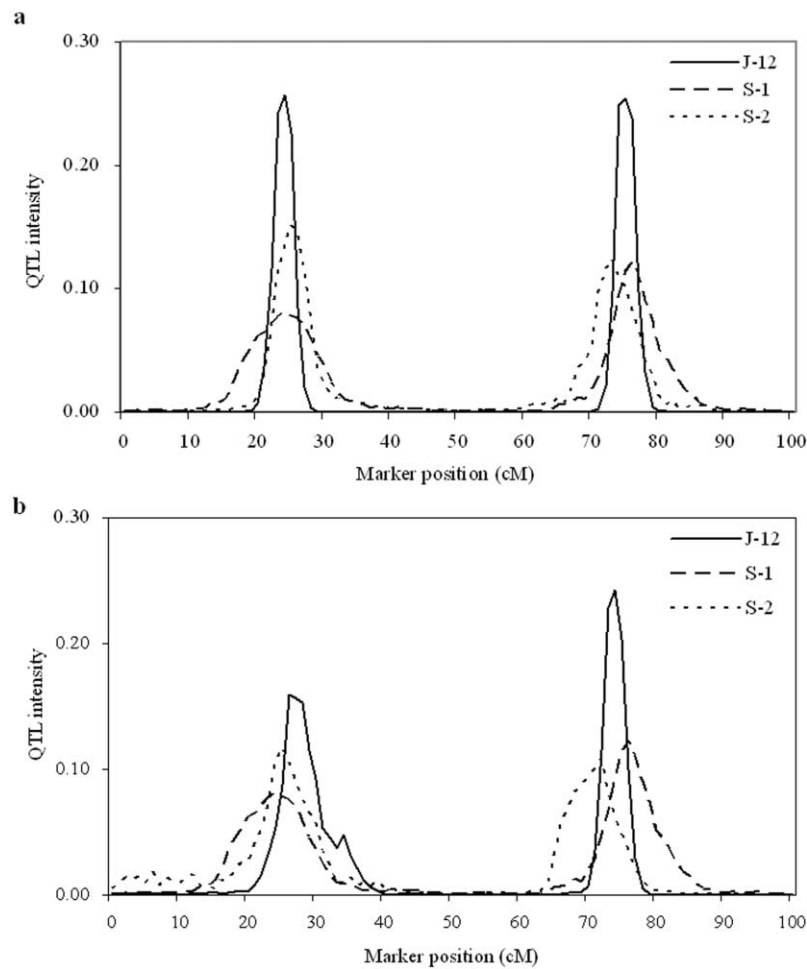
Figure 1. Approximate posterior distributions of QTL locations based on QTL intensity with step length of 1 cM from joint mapping and single-trait mapping in both simulation experiments. *a,* Experiment 1 with two continuous traits. *b,* Experiment 2 with a mixture of one quantitative and one binary trait. J-12, S-1, and S-2 denote joint mapping for both traits and separate mapping for trait 1 and trait 2, respectively. The true number of QTLs is two, located at positions 25 and 75 cM.

Table 2. Bayesian Estimates of Joint Mapping and Single-Trait Mapping from the First Simulation Experiment

Analysis Method[a] and Interval (~cM)	Sum of the QTL Intensity	QTL Location (cM)	$\sigma^2_{qj,1}$	$\sigma^2_{qj,2}$	$\sigma_{qj,12}$	$\sigma^2_{a,1}$	$\sigma^2_{a,2}$	$\sigma_{a,12}$	$\sigma^2_{e,1}$	$\sigma^2_{e,2}$	$\sigma_{e,12}$	μ
J-12:						.5512 (.0146)	.3669 (.0180)	.1793 (.0101)	.9946 (.0095)	1.0704 (.0163)	.0590 (.0058)	
20–28	.9940	24.3411 (2.2529)	.2138 (.0103)	.3170 (.0123)	−.0135 (.0019)							.0376$^{\mu_1}$ (.0019)
72–81	.9976	75.5649 (2.0745)	.2288 (.0146)	.2947 (.0132)	.2360 (.0071)							.0251$^{\mu_2}$ (.0016)
S-1:						.8307 (.1394)			.8992 (.0270)			.0401$^{\mu_1}$ (.0044)
18–33	.8924	25.0136 (14.7105)	.1438 (.0304)									
71–84	.9260	77.1874 (14.3854)	.1869 (.0322)									
S-2:							.5325 (.0624)			.9813 (.0276)		.0286$^{\mu_2}$ (.0030)
21–32	.9168	25.7168 (6.5950)		.2492 (.0336)								
68–81	.9364	73.9050 (9.5818)		.2660 (.0296)								

NOTE.—Both traits $\mathbf{y}_1$ and $\mathbf{y}_2$ are continuous. Posterior mean squared errors of the estimates are given in parentheses. $\sigma^2_{qj,1}$, $\sigma^2_{qj,2}$, and $\sigma_{qj,12}$ are the jth ($j = 1,2$) QTL (co)variances of the two traits; $\sigma^2_{a,1}$, $\sigma^2_{a,2}$, and $\sigma_{a,12}$ are the polygenic (co)variances of the two traits; and $\sigma^2_{e,1}$, $\sigma^2_{e,2}$, and $\sigma_{e,12}$ are the residual (co)variances of the two traits.

[a] J-12 = joint mapping for traits $\mathbf{y}_1$ and $\mathbf{y}_2$; S-1 = separate mapping for trait $\mathbf{y}_1$; S-2 = separate mapping for trait $\mathbf{y}_2$.

Table 3. Bayesian Estimates of Joint Mapping and Single-Trait Mapping from the Second Simulation Experiment

Analysis Method[a] and Interval (~cM)	Sum of the QTL Intensity	QTL Location (cM)	$\sigma^2_{qj,1}$	$\sigma^2_{qj,2}$	$\sigma_{qj,12}$	$\sigma^2_{a,1}$	$\sigma^2_{a,2}$	$\sigma_{a,12}$	$\sigma^2_{e,1}$	$\sigma_{e,12}$	μ
J-12:						.7806 (.0965)	.5270 (.0505)				
23–33	.9848	26.3234 (5.9781)	.2364 (.0203)	.3389 (.0333)	−.0026 (.0040)			.2156 (.0111)	.9000 (.0211)	−.0128 (.0056)	−.0394$^{\mu_1}$ (.0038)
69–79	.9658	74.4096 (2.3955)	.2103 (.0246)	.3472 (.0369)	.2277 (.0123)						−.0291$^{\mu_2}$ (.0029)
S-1:											
18–33	.8924	25.0136 (14.7105)	.1438 (.0304)								
71–84	.9260	77.1874 (14.3854)	.1869 (.0322)			.8307 (.1394)			.8992 (.0270)		.0401$^{\mu_1}$ (.0044)
S-2:							.6591 (.1982)				−.0343$^{\mu_2}$ (.0050)
17–32	.9214	25.1521 (13.1565)		.3573 (.0595)							
65–78	.8853	72.1356 (21.4333)		.4044 (.0916)							

NOTE.—The continuous trait $\mathbf{y}_1$ and the binary trait $\mathbf{w}_2$ are considered in the analysis. Posterior mean squared errors of the estimates are given in parentheses. $\sigma^2_{qj,1}$, $\sigma^2_{qj,2}$, and $\sigma_{qj,12}$ are the jth ($j = 1,2$) QTL (co)variances of the two traits; $\sigma^2_{a,1}$, $\sigma^2_{a,2}$, and $\sigma_{a,12}$ are the polygenic (co)variances of the two traits; and $\sigma^2_{e,1}$, $\sigma^2_{e,2}$, and $\sigma_{e,12}$ are the residual (co)variances of the two traits.

[a] J-12 = joint mapping for traits $\mathbf{y}_1$ and $\mathbf{y}_2$; S-1 = separate mapping for trait $\mathbf{y}_1$; S-2 = separate mapping for trait $\mathbf{w}_2$.

QTLs with pleiotropic effects, the corresponding QTL co-variance estimates accordingly suggest a high QTL-effect correlation (close to 1) between the two traits. The results presented here show that the genetic mechanism of QTLs can be preliminarily probed through the respective estimates of covariance. Moreover, appropriate statistical testing—for example, use of the BF—can be used to further discriminate coincident linkage from pleiotropy.

The estimates of the overall mean, the polygene (co)variances, and the residual (co)variances are also given in tables 2 and 3 for the two simulations. It is clear that the covariance estimates for polygene and residual effects between traits can be obtained only from joint analyses. It appears that the estimates of these parameters are largely around the simulated values. However, the estimates from joint analyses are much closer to the true parameters and have smaller posterior mean squared errors than do those from separate analyses in both experiments.

The results presented in tables 1–3 and figure 1 demonstrate that the developed multivariate version of Bayesian mapping performs well for multiple continuous traits as well as for a mixture of traits. Compared with separate analyses, our joint-mapping method can enhance the precision and power in QTL detection as well as in estimation of other parameters. Moreover, from simulation results, it appears that joint mapping for multiple continuous traits always outperforms that for a mixture of traits, probably because of information loss in binary phenotypic data.

In table 4, we present the estimates of the BF for testing pleiotropy versus coincident linkage in the two simulation experiments. We calculated the BF on the basis of the MCMC process, to test coincident linkage of two nonpleiotropic QTLs against pleiotropy of a common QTL. The test was performed for each interval region where the QTL had been found in the initial joint analysis. For the first experiment, two QTL intervals, one 20–28 cM and the other 72–81 cM, were considered. For the second experiment, the tests were conducted in the two regions, one 23–33 cM and the other 69–79 cM. According to the commonly used criteria for choice of model based on the BF,[22,40] both estimates of the BF for the first set of QTLs in the two simulation experiments are far less than 0.1, which indicates a much higher posterior probability for the coincident linkage model and provides strong evidence for accepting the hypothesis of coincident linkage of two nonpleiotropic QTLs within this interval. For the second set of QTLs, we obtain the BF values from both simulations, which are much larger than 10.0. This gives strong evidence favoring the hypothesis of the second set of QTLs having pleiotropic effects on both traits. On the basis of the results in table 4, we accept the hypothesis that there exists coincident linkage of two nonpleiotropic QTLs within the interval of the first set of QTLs and one QTL with pleiotropic effects within the interval of the second set of QTLs for both simulation experiments. These results are consistent with the simulated genetic mechanism for the respective QTLs.

Table 4. BF for Testing Pleiotropy and Coincident Linkage for the Two QTLs Detected by Joint Mapping in Both Experiments

Experiment and Detected QTL	Interval (cM)	BF ($Model_{plei.}$ vs. $Model_{link.}$)
1[a]:		
QTL 1	20–28	3.55×10^{-3}
QTL 2	72–81	1.87×10^{2}
2[b]:		
QTL 1	23–33	8.26×10^{-2}
QTL 2	69–79	$.49 \times 10^{2}$

[a] Experiment 1 considered two continuous traits.
[b] Experiment 2 considered one continuous trait and one binary trait.
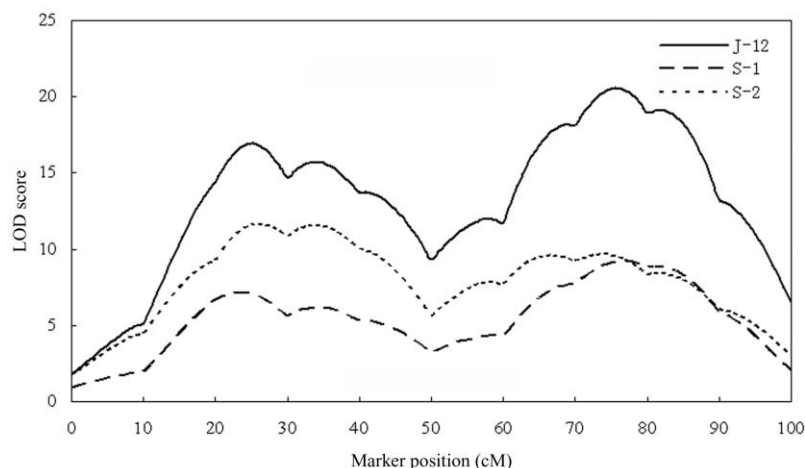
To confirm our proposed joint-mapping method, the multiple continuous data set simulated in the first experiment was also analyzed using the program SOLAR[42] version 3.04. Figure 2 shows the LOD-score profiles for J-12, S-1, and S-2 along the chromosome. We can see that each LOD-score profile shows two peaks. These two peaks are at positions 25 cM and 76 cM for J-12, 24 cM and 77 cM for S-1, and 26 cM and 74 cM for S-2, all overlapping the true locations of the two simulated QTLs. Comparing the shapes of different LOD-score profiles, we found that the LOD-score profile for J-12 has the highest peaks among the three types of analyses. These results are consistent with those obtained using our Bayesian joint-mapping method. Finally, in the J-12 analysis, we ran SOLAR to test pleiotropy versus coincident linkage at the two peak positions. At position 25 cM, we were able to reject the hypothesis of complete pleiotropy with an extremely small P value ($P_1 = 6.91 \times 10^{-9}$) while accepting the hypothesis of coincident linkage ($P_0 = .2305$). At position 76 cM, we rejected the hypotheses of coincident linkage with a very small P value ($P_0 = 5.87 \times 10^{-10}$) but accepted the hypothesis of complete pleiotropy ($P_1 = .3741$). These results are also in agreement with those obtained by using our testing method based on the BF.

It is apparent from figures 1 and 2 that the QTL intensity curves are much narrower and the peaks are more precisely located close to the simulated true QTLs with our method than with SOLAR. In addition, to draw generalizations from the studies, we simply simulated 50 independent samples because of the high computational demands of the MCMC algorithm for simulation 1, on the basis of the same parameter settings. Although 50 replicates are not sufficient to accurately evaluate the statistical power, a general idea about the performance of our method can be obtained in this way. In each replicate, we claimed a QTL as "detected" if there was a remarkable QTL intensity peak around the "true" position. We present the average estimates for parameters of each QTL obtained from 50 replicates in all analyses. As presented in table 5, higher powers and fewer posterior mean squared errors were largely seen in the joint analyses.

Real-Data Analyses

Bone mineral density (BMD) is the primary predictor of bone strength and fracture risk.[43–45] Areal bone size (ABS) is also a significant determinator of bone strength and has shown independence in predicting fracture risk.[46–49] Both BMD and ABS have strong determination, with heritabilities >50% in human populations.[43–45] Studies show that BMD and ABS are genetically correlated and may share some genetic determination.[49]

To identify genomic regions that harbor QTLs having common effects on BMD and ABS at the forearm, we recently performed a bivariate genomewide linkage scan in a large sample of 451 white pedigrees comprising 4,498 individuals (X. Liu, Y. Liu, J. Liu, Y. Pei, P. Xiao, H. Shen, J. Deng, R. R. Recker, H.-W. Deng, unpublished data). Our preliminary results demonstrated a suggestive signal at chromosome 7p15, which showed coincident linkage. To further validate our proposed method, we performed a bivariate linkage scan, using our method on the same real-data set. Here, we performed linkage analyses only on chromosome 7 (encompassing 21 genotyped microsatellite markers), where the highest linkage peak was observed. For ease of calculation, we picked 364 independent nuclear families, comprising 2,007 subjects, from the original sample set. The number of sibs ranges from 2 to 14 across the families, with an average of 5.5.

As we did in simulation studies, we conducted Bayesian mapping for the real-data set, using both univariate and bivariate analyses. The QTL intensity distributions obtained from joint analysis (for BMD and ABS) and from separate analyses for BMD and ABS are given in figure 3. A QTL intensity peak was obtained in each analysis (joint analysis and separate analyses), highly favoring a single-QTL model. Furthermore, the joint analysis of BMD and ABS suggested a chromosome region having common effects on both traits. However, as clearly shown in figure

3, it is difficult to identify the linkage to ABS by performing single-trait analysis, because of the modest genetic effects and, thus, the weak linkage signals. As expected, we benefited from joint analysis under this situation because it yielded a more intense peak than did separate analyses for BMD and ABS. To further differentiate pleiotropy versus coincident linkage, we calculated the BF on the chromosomal interval 43–44 cM, where the QTL intensity peak was obtained. A coincident linkage model was favored from the calculation of the BF (0.0891). Our findings herein from the use of the Bayes method are largely consistent with those from our earlier linkage studies that used variance-components analyses implemented in SOLAR.

We also obtained the estimates of the number of QTLs and other relative parameter estimates for both BMD and

Table 5. Average Estimates of QTL Parameters Obtained from 50 Replicates in Each Analysis of the First Simulation Study

QTL and Analysis Method	Parameter Estimate		
	Position (cM)	V_{qtl}	Power
1:			
J-12:	24.5072 (2.6319)		1.00 (50/50)
Trait 1		.2307 (.0252)	
Trait 2		.2856 (.0183)	
S-1	26.1832 (9.1047)	.1659 (.0344)	.92 (46/50)
S-2	25.4635 (7.2281)	.2706 (.0375)	.98 (49/50)
2:			
J-12:	75.6623 (2.7051)		1.00 (50/50)
Trait 1		.2086 (.0173)	
Trait 2		.3120 (.0198)	
S-1	75.3598 (8.7454)	.1795 (.0426)	.94 (47/50)
S-2	74.0064 (7.3095)	.2941 (.0373)	1.00 (50/50)

NOTE.—The averaged posterior mean squared errors over 50 replicates are given in parentheses.

Figure 2. LOD-score profiles of QTL mapping from joint analysis of two continuous traits (J-12) and separate analysis for trait 1 (S-1) and trait 2 (S-2), with the use of SOLAR software. Two QTLs are simulated at positions 25 and 75 cM.
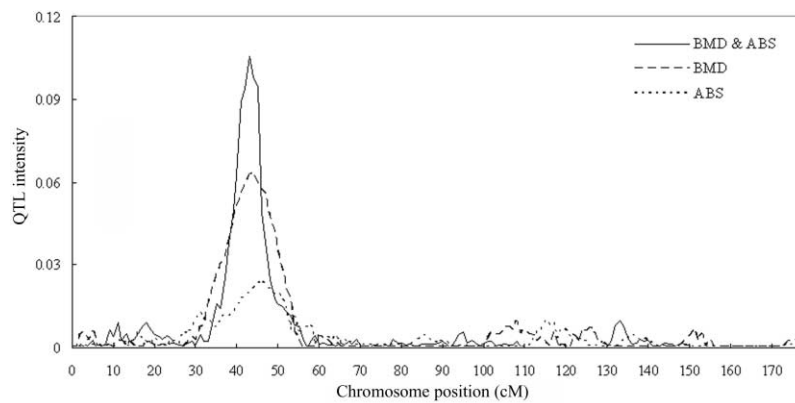
Figure 3. QTL intensity distributions based on joint analysis and separate analyses for BMD and ABS of wrist on chromosome 7

ABS (data not shown). Compared with separate single-trait analyses, the joint analysis showed less estimation variance for the QTL number, higher posterior probability for the interval of identified QTLs, and lower SEs of parameter estimation. The increased estimation accuracy of parameters and higher posterior QTL intensity in real-data analysis clearly indicate that Bayesian joint analysis renders advantages over separate analyses, which provides further evidence to support our findings demonstrated in simulation studies.

Discussion

In this study, a joint-mapping method was developed using the MCMC algorithm in a Bayesian framework. The proposed method can handle multiple continuous traits, as well as a mixture of continuous and discrete traits (such as disease status).

Our proposed method has significant practical importance for genetic mapping of human complex diseases, such as essential hypertension (MIM 145500), type 2 diabetes (MIM 125853), osteoporosis (MIM 166710), obesity (MIM 601665), etc. When we map disease-susceptibility genes, the disease status is recorded and is generally treated as a dichotomous trait in statistical analysis. Moreover, in practical studies, measures of some highly related quantitative traits (risk factors of the disease) are often measured. For example, BMD, a quantitative measure of bone mass, is often used to evaluate the risk of osteoporotic fracture. Joint mapping that uses information from various data types can essentially increase precision of parameter estimations and, thus, the chance of identifying disease-susceptibility genes. Most importantly, given that quantitative risk factors may not share 100% of genetic determination with disease outcomes,[50,51] joint analyses of disease and quantitative risk factors can identify those genes that are shared and that are distinct between these two different types of disease-related phenotypes.

We performed an MCMC process to generate the mar-

ginal posterior distribution of the parameters of interest, which is not used in other existing joint-mapping methodologies. Inference for all parameters is based on their respective posterior samples. Two types of simulation experiments were conducted to evaluate the performance of the proposed method. The results clearly demonstrate that, compared with the single-trait analysis, our method can identify QTLs with higher posterior probabilities and narrower intervals and can estimate unknowns with smaller posterior mean squared errors. These findings suggest that the Bayesian joint mapping introduced here can improve the precision of parameter estimation and produces stronger evidence to identify QTLs.

We adopted a Bayesian method and implemented the procedure, using the MCMC algorithm. This is different from other existing joint-mapping methods. We used the Bayesian MCMC approach because of the following obvious advantages in comparison with the commonly used likelihood-ratio test methods: (1) One problem with the traditional likelihood-ratio test is that it is not always easy to obtain maximum-likelihood estimators either by maximizing the likelihood or by using an iterative algorithm, such as expectation maximum. But, in a Bayesian framework, we can avoid making such an optimization for joint-mapping analysis. (2) The Bayesian MCMC approach allows us to directly measure the probability as the evidence of the identified QTL and respective interval, which makes it easier to compare our results with the results of other experiments. To our knowledge, no other existing statistical paradigm is able to do this. In contrast, the *P* values obtained from likelihood-ratio tests cannot be readily compared among different experiments. Furthermore, we often encounter the multiple-comparison problem in statistical tests, and it is still challenging to deal with several dependent tests. Bayesian inference has an obvious advantage over traditional frequentist approaches, in terms of multiple-comparison testing, in that degree of belief is quantified by the posterior probabilities. Therefore, one can avoid, to some extent, the potential illogical conclu-

sions that arise from an simple "accept/reject" decision process.[52] (3) In a Bayesian framework, estimates of all unknowns are obtained from their posterior distributions, and, thus, we can avoid the use of asymptotic approximation by employing Fisher information. (4) The Bayesian method can construct more-complicated models by incorporating multiple levels of randomness and combining information from different sources.

Joint mapping offers an opportunity to distinguish coincident linkage from pleiotropy. In this study, we successfully distinguished the coincident linkage model and the pleiotropic model by employing a BF method for Bayesian joint mapping. To confirm the results of our BF method, we used SOLAR to test pleiotropy versus coincident linkage on the basis of a likelihood-ratio test[6] for two continuous traits that are simulated in the first experiment. The results from SOLAR are highly in agreement with those from our BF method. However, our method gives much better precision and resolution, as shown in figures 1 and 2. So far, a suite of methods have been developed for joint mapping, most of which[1,5–7,10,53] were based on likelihood-ratio tests. Departing from those likelihood-ratio–based methods, we used the BF to distinguish coincident linkage from pleiotropic models. The BF automatically generates the posterior probabilities for each model. The major advantage of the BF method is that it does not rely on any asymptotic results, which is a must for the likelihood-based methods, and it can provide exact outcomes with any amount of information. The more information used, the higher the probability that the "best" model will be obtained.[54] Recently, Varona et al.[40] proposed a method that uses the BF to distinguish between linked and pleiotropic QTLs under the linear regression model. In their approach, they adopted a Gibbs sampler to compute the posterior distribution of the QTL location under the linkage model. However, it is not clear if this approach could be adapted to identify the coincidently linked QTLs residing in the same marker bracket.

In the MCMC process, efficiency of the two samplers—that is, the Gibbs sampler and RJ-MCMC—depends on the mixing property of the Markov chain, which in turn is determined by the sampling scheme adopted and the parameterization involved in the model. To ensure that the Markov chain can mix well, we improved our sampling scheme when conducting RJ-MCMC for joint mapping in both simulation experiments. Specifically, when determining the acceptance probability of adding a new QTL via RJ-MCMC, we used the Gibbs sampler to perform an extra 10 updates of parameters for the proposed QTL, which greatly accelerated the convergence of the Markov chain, as expected.

In our studies, we adopted a conservative approach to deal with the convergence of MCMC. That is, thinning and burn-in values for all MCMC analyses were set to 50 and 50,000 (1,000 × 50), respectively. Empirically, these values are much higher than what are probably needed to ensure convergence of MCMC. Furthermore, we ran

CODA, a widely used R package, to perform formal convergence tests, using the method of Raftery and Lewis.[55] From the diagnostic results for the two simulations (data not shown), it is clear that dependence factors for all parameters are <5.0, which indicates that the thinning and burn-in values adopted herein are reasonable for ensuring the convergence of MCMC. As for RJ-MCMC output for the QTL parameters, we generated a time-series plot of the sampled points for the number of QTLs in each analysis rather than use a formal convergence test. This is a common practice in Bayesian mapping.[16,17,28,29] It also shows that the RJ-MCMC algorithm mixes well over the number of QTLs. In addition, we calculated autocorrelation coefficients for samples of each parameter by running CODA. By comparing the amount of autocorrelation for each chain of parameters between single-trait analyses and joint analyses, we found that higher autocorrelation often occurred in joint analyses. This further suggests that we should generate a longer chain of parameters to reach convergence with the increasing of the number of traits involved in joint analyses, although the length of 1×10^5 is large enough in our studies.

In the RJ-MCMC process, the information about a QTL is totally lost when we delete this QTL from the model, which could influence the acceptance probability and the mixing behavior.[56] Recently, two promising methods proposed for addressing the problem of RJ-MCMC—a stochastic search variable selection (SSVS) approach[56,57] and a Bayesian shrinkage method[58,59]—were proposed for Bayesian mapping in experimental populations. In the SSVS procedure, a mixture prior is adopted to explicitly make a probabilistic statement about the inclusion of a QTL, and the markers with significant effects can be identified as those with higher posterior probability included in the model. As for Bayesian shrinkage analysis, each marker or marker interval is assumed to be associated with a QTL. If a marker or marker interval is not associated with any QTL, the corresponding QTL effect will be forced to shrink toward zero. Both methods above can largely avoid the problems seen in RJ-MCMC. Given the potential advantages of these two methods over RJ-MCMC, we will pursue further studies to extend them to a random model, which involves QTLs having infinite alleles. Under such a model, each QTL has infinite genotypes, and the QTL effects may vary across different individuals. Hence, with shrinkage analysis, it is impossible to shrink the designated QTL effect to zero if there is no associated QTL. Similarly, in SSVS, it is hard to make a probabilistic statement about the inclusion of a QTL according to the QTL effect. As such, with variance-components analyses, it may be a natural way to perform Bayesian shrinkage and SSVS analyses based on QTL variances instead of QTL effects. To develop such an extension of the two methods to an infinite-allele model, substantial effort is still needed.

The presented method is a variance-components joint-mapping procedure, which means that the infinite-allele model is used and the (co)variances of QTLs and other

random effects in the model are treated as parameters of interest. Our method is quite suitable for natural outbred populations, such as humans. This is because the number of alleles at the locus is generally unknown in natural populations. Although the biallelic model could be appropriate in most situations, it may cause potential bias in estimation if the true number of the QTL alleles exceeds two. Using the variance-components model, we only need to estimate the (co)variances of respective QTLs, regardless of the number of the QTL alleles, which substantially improves the robustness of the Bayesian joint-mapping method.

In the simulation studies, for the purpose of simplicity, we used only full-sib families to show the performance of our proposed method. It is straightforward to extend our method to complex pedigrees by using a general approach to calculate the IBD matrix related to each QTL. Such a general approach has been developed by many investigators.[42,60,61] In addition, although our proposed method is focused on linkage analyses, it can be naturally extended to multiple-trait fine mapping by using both linkage analysis information and linkage disequilibrium (LD) information if highly dense markers are studied. Generally, combining LD information and linkage analysis information can significantly improve the power and precision of QTL mapping. For this situation, the extra effort needed is to remove the original assumption that founder QTL alleles are unrelated, and IBD probabilities between the founder QTL alleles can be estimated through the existing well-developed methods.[62,63]

In our analyses, we assumed no correlation between the residuals. However, we acknowledge that the residual correlation may have significant influence on QTL mapping in joint analysis. From the findings of Jiang and Zeng[1] and Allison et al.[64] in joint-mapping analysis, the greatest power increases usually occur when the QTL-induced correlation is opposite in sign to the correlation induced by residual factors. Our proposed method can directly handle multiple phenotypic data without any modification when correlated residuals are assumed. We will investigate the efficiencies of our Bayesian joint-mapping procedure under several scenarios of residual correlations in our future endeavors.

Acknowledgments

Appendix A

Joint Analysis with Multiple Continuous Traits and Multiple Ordered Categorical Traits

Assume m_1 normal continuous traits and m_2 ordered categorical traits, each having $C_s(s = 1, \ldots, m_2)$ ordered categories delimited by $C_s - 1$ thresholds ($m = m_1 + m_2$). We denote $\mathbf{y} = (\mathbf{y}'_1, \ldots, \mathbf{y}'_k, \ldots, \mathbf{y}'_n)'$ as an $nm_1 \times 1$ vector of observations with n individuals on m_1 continuous traits and $\mathbf{w} = (\mathbf{w}'_1, \ldots, \mathbf{w}'_k, \ldots, \mathbf{w}'_n)'$ and $\xi = (\xi'_1, \ldots, \xi'_k, \ldots, \xi'_n)'$ as $nm_2 \times 1$ vectors of phenotypes and the unobservable liability of n individuals on m_2 discrete traits, respectively. We also denote $t_{c,s}$ as the cth threshold for the sth discrete trait ($c = 1, \ldots, C_s - 1$). Similarly, if liability vector ξ can be drawn from its fully conditional distribution in each iteration of the MCMC process, Bayesian mapping for mixed types of $m_1 + m_2$ traits can be performed using the method for m continuous traits detailed above. Since discrete data considered here fall in more than two categories, we cannot predetermine the threshold value $t_{c,s}$ as done for binary traits. Thus, no constraint is superimposed to the (co)variance matrix of residual effects among m_2 discrete traits. For convenience of explanation, we present the $m \times m$ (co)variance matrix $\mathbf{R}$ among all $m_1 + m_2$ traits as

$$\mathbf{R} = \left(\begin{array}{ccc|ccc} \sigma^2_{y_1} & \cdots & \sigma_{y_1,y_{m1}} & \sigma_{y_1,\xi_1} & \cdots & \sigma_{y_1,\xi_{m2}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{y_{m1},y_1} & \cdots & \sigma^2_{y_{m1}} & \sigma_{y_{m1},\xi_1} & \cdots & \sigma_{y_{m1},\xi_{m2}} \\ \hline \sigma_{\xi_1,y_1} & \cdots & \sigma_{\xi_1,y_{m1}} & \sigma^2_{\xi_1} & \cdots & \sigma_{\xi_1,\xi_{m2}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{\xi_{m2},y_1} & \cdots & \sigma_{\xi_{m2},y_{m1}} & \sigma_{\xi_{m2},\xi_1} & \cdots & \sigma^2_{\xi_{m2}} \end{array} \right)$$

$$= \begin{pmatrix} \Omega_y & \Omega_{y,\xi} \\ \Omega_{\xi,y} & \Omega_\xi \end{pmatrix} \,,$$

where $m \times m$ matrix $\mathbf{R}$ is partitioned into four matrices—that is, an $m_1 \times m_1$ residual (co)variance matrix Ω_y among traits $\mathbf{Y}$, an $m_2 \times m_2$ residual (co)variance matrix Ω_ξ among traits ξ, and an $m_1 \times m_2$ residual covariance matrix $\Omega_{y,\xi}$ between traits Y and ξ and its transpose, $\Omega_{\xi,y}$.

According to the assumption of the threshold model,[65] it is straightforward to derive the fully conditional distribution of liability ξ_k for individual k:

$$p(\xi_k | l, \lambda, \theta, \mathbf{w}, \xi_{-k}, \mathbf{y}) \propto p(\xi_k, \mathbf{y}_k | l, \lambda, \theta) p(\mathbf{w}_k | \xi_k)$$

$$\propto p(\xi_k | \mathbf{y}_k, l, \lambda, \theta) \prod_{s=1}^{m_2} \left[\sum_{c=1}^{C_s} I(t_{c-1,s} < \xi_{k,s} \leq t_{c,s}) I(w_{k,s} = c) \right] , \quad \text{(A1)}$$

where $\xi_{k,s}$ and $w_{k,s}$ are the sth elements in vectors ξ_k and $\mathbf{w}_k$, respectively, for the sth discrete trait. $w_{k,s} = c$ means that the observation on the sth discrete trait of individual k falls in category c. Specifically, $t_{0,s} = -\infty$, $t_{Cs,s} = +\infty$, and $t_{0,s} < t_{1,s} < \ldots < t_{Cs,s}$ for all s. In equation (A1), the fully conditional distribution of liability vector ξ_k is in the form of

a truncated, conditional multinormal distribution; each element in ξ_k corresponding to categorical trait s is truncated at thresholds $t_{c-1,s}$ and $t_{c,s}$. The mean vector of the truncated conditional multinormal distribution is given as

$$\mathrm{E}(\xi_k | l,\boldsymbol{\lambda},\boldsymbol{\theta},\mathbf{w},\xi_{-k},\mathbf{y}) = \boldsymbol{\beta}_{k,\xi} + \mathbf{u}_{k,\xi} + \sum_{i=1}^{l} \mathbf{q}_{i,(k,\xi)}$$

$$+\Omega_{\xi,y}(\Omega_y)^{-1}[\mathbf{y}_k - \beta_{k,y} - \mathbf{u}_{k,y} - \sum_{i=1}^{l} \mathbf{q}_{i,(k,y)}] \ ,$$

and the (co)variance matrix for the distribution is

$$\mathrm{Var}\,(\xi_k | l,\boldsymbol{\lambda},\boldsymbol{\theta},\mathbf{w},\xi_{-k},\mathbf{y}) = \Omega_\xi - \Omega_{\xi,y}(\Omega_y)^{-1}\Omega_{y,\xi} \ .$$

Since we cannot predefine the thresholds for each categorical trait, the Gibbs sampler is also employed to generate those thresholds in the MCMC process, simply on the basis of a uniform distribution:

$$p(t_{c,s}|\xi_{.,s},\mathbf{w}_{.,s},t_{c-1,s},t_{c+1,s}) =$$

$$\frac{1}{\min\,[\min\,(\xi_{.,s}\,|\,\mathbf{w}_{.,s}=c+1),t_{c+1,s}] - \max\,[\max\,(\xi_{.,s}\,|\,\mathbf{w}_{.,s}=c),t_{c-1,s}]} \ ,$$

$$(A2)$$

where $\xi_{.,s}$ and $\mathbf{w}_{.,s}$ denote subvectors of ξ and $\mathbf{w}$ related to categorical trait s. For each trait $s(s = 1,\dots,m_2)$, we successively draw the samples of all thresholds $t_{1,s},\dots,t_{Cs-1,s}$ in turn via equation (A2).

Compared with the original MCMC process for joint analyses with multiple continuous traits, additional efforts are needed for mixed types of traits considered here, including the following:

1. In step 1, we should assign proper starting values to thresholds of each categorical trait according to respective categories and then generate ξ in terms of threshold values $t_{c,s}$ and $\mathbf{w}$.
2. In step 2, the samples of $t_{c,s}$ are drawn from equation (A2), followed by performance of the Gibbs sampler, to sample ξ on the basis of equation (A1). After that, the joint analysis with m_1 continuous traits and m_2 liability variables is essentially the same as that for m continuous traits.

Web Resources

The URLs for data presented herein are as follows:

Bayes Mapping for Multiple Traits, http://l.web.umkc.edu/liujian/ (for the program for our proposed method)
CODA, http://cran.us.r-project.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for essential hypertension, type 2 diabetes, osteoporosis, and obesity)
SOLAR, http://www.sfbr.org/solar/

References

1. Jiang C, Zeng ZB (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111–1127
2. Weller JI, Wiggans GR, Vanraden PM, Ron M (1996) Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multitrait experiment. Theor Appl Genet 92:998–1002
3. Lange C, Whittaker JC (2001) Mapping quantitative trait loci using generalized estimating equations. Genetics 159:1325–1337
4. Huang J, Jiang Y (2003) Genetic linkage analysis of a dichotomous trait incorporating a tightly linked quantitative trait in affected sib pairs. Am J Hum Genet 72:949–960
5. Lund MS, Sorensen P, Guldbrandtsen B, Sorensen DA (2003) Multitrait fine mapping of quantitative trait loci using combined linkage disequilibria and linkage analysis. Genetics 163:405–410
6. Almasy L, Dyer TD, Blangero J (1997) Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. Genet Epidemiol 14:953–958
7. Knott SA, Haley CS (2000) Multitrait least squares for quantitative trait loci detection. Genetics 156:899–911
8. Xu C, Li Z, Xu S (2005) Joint mapping of quantitative trait loci for multiple binary characters. Genetics 169:1045–1059
9. Williams JT, Begleiter H, Porjesz B, Edenberg HJ, Foroud T, Reich T, Goate A, Van Eerdewegh P, Almasy L, Blangero J (1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. Am J Hum Genet 65:1148–1160
10. Williams JT, Van Eerdewegh P, Almasy L, Blangero J (1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. I. Likelihood formulation and simulation results. Am J Hum Genet 65:1134–1147
11. Meuwissen TH, Goddard ME (2004) Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genet Sel Evol 36:261–279
12. Hackett CA, Meyer RC, Thomas WT (2001) Multi-trait QTL mapping in barley using multivariate regression. Genet Res 77:95–106
13. Henshall JM, Goddard ME (1999) Multiple-trait mapping of quantitative trait loci after selective genotyping using logistic regression. Genetics 151:885–894
14. Mangin B, Thoquet P, Grimsley N (1998) Pleiotropic QTL analysis. Biometrics 54:88–99
15. Gilbert H, Le Roy P (2003) Comparison of three multitrait methods for QTL detection. Genet Sel Evol 35:281–304
16. Yi N, Xu S (2000) Bayesian mapping of quantitative trait loci under the identity-by-descent-based variance component model. Genetics 156:411–422
17. Yi N, Xu S (2000) Bayesian mapping of quantitative trait loci for complex binary traits. Genetics 155:1391–1403
18. Dunson DB (2001) Commentary: practical advantages of Bayesian analysis of epidemiologic data. Am J Epidemiol 153:1222–1226
19. Yi N, Xu S, George V, Allison DB (2004) Mapping multiple quantitative trait loci for ordinal traits. Behav Genet 34:3–15
20. Yi N, Xu S, Allison DB (2003) Bayesian model choice and search strategies for mapping interacting quantitative trait loci. Genetics 165:867–883
21. Yi N, Xu S (2001) Bayesian mapping of quantitative trait loci under complicated mating designs. Genetics 157:1759–1771

22. Satagopan JM, Yandell BS, Newton MA, Osborn TC (1996) A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. Genetics 144:805–816

23. von Rohr P, Hoeschele I (2002) Bayesian QTL mapping using skewed Student-t distributions. Genet Sel Evol 34:1–21

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

25. Uimari P, Thaller G, Hoeschele I (1996) The use of multiple markers in a Bayesian method for mapping quantitative trait loci. Genetics 143:1831–1842

26. Hoeschele I, Vanranden PM (1993) Bayesian analysis of linkage between genetic markers and quantitative trait loci. I. Prior knowledge. Theor Appl Genet 85:953–960

27. Hoeschele I, Vanranden PM (1993) Bayesian analysis of linkage between genetic markers and quantitative trait loci. II. Combining prior knowledge with experimental evidence. Theor Appl Genet 85:946–952

28. Sillanpaa MJ, Arjas E (1999) Bayesian mapping of multiple quantitative trait loci from incomplete outbred offspring data. Genetics 151:1605–1619

29. Sillanpaa MJ, Arjas E (1998) Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics 148:1373–1388

30. Chan JSK, Kuk AYC (1997) Maximum likelihood estimation for probit-linear mixed models with correlated random effects. Biometrics 53:86–97

31. Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82:711–732

32. Xu S, Gessler DD (1998) Multipoint genetic mapping of quantitative trait loci using a variable number of sibs per family. Genet Res 71:73–83

33. Sorensen D, Gianola D (2002) Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer-Verlag, New York

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

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

36. Gelman A, Carlin JB, Stern HS, Rubin DB (1995) Bayesian data analysis. Chapman and Hall, London

37. Jannink JL, Fernando RL (2004) On the Metropolis-Hastings acceptance probability to add or drop a quantitative trait locus in Markov chain Monte Carlo-based Bayesian analyses. Genetics 166:641–643

38. Sillanpaa MJ, Gasbarra D, Arjas E (2004) Comment on "On the Metropolis-Hastings acceptance probability to add or drop a quantitative trait locus in Markov chain Monte Carlo-based Bayesian analyses." Genetics 167:1037

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

40. Varona L, Gomez-Raya L, Rauw WM, Clop A, Ovilo C, Noguera JL (2004) Derivation of a Bayes factor to distinguish between linked or pleiotropic quantitative trait loci. Genetics 166:1025–1035

41. Bulmer MG (1980) The mathematical theory of quantitative genetics. Oxford University Press, New York

42. Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62: 1198–1211

43. Liu YJ, Shen H, Xiao P, Xiong DH, Li LH, Recker RR, Deng HW (2006) Molecular genetic studies of gene identification for osteoporosis: a 2004 update. J Bone Miner Res 21:1511–1535

44. Liu YZ, Liu YJ, Recker RR, Deng HW (2003) Molecular studies of identification of genes for osteoporosis: the 2002 update. J Endocrinol 177:147–196

45. Ralston SH, de Crombrugghe B (2006) Genetic regulation of bone mass and susceptibility to osteoporosis. Genes Dev 20: 2492–2506

46. Augat P, Reeb H, Claes LE (1996) Prediction of fracture load at different skeletal sites by geometric properties of the cortical shell. J Bone Miner Res 11:1356–1363

47. Duan Y, Parfitt A, Seeman E (1999) Vertebral bone mass, size, and volumetric density in women with spinal fractures. J Bone Miner Res 14:1796–1802

48. Seeman E, Duan Y, Fong C, Edmonds J (2001) Fracture site-specific deficits in bone size and volumetric density in men with spine or hip fractures. J Bone Miner Res 16:120–127

49. Deng HW, Xu FH, Davies KM, Heaney R, Recker RR (2002) Differences in bone mineral density, bone mineral content, and bone areal size in fracturing and non-fracturing women, and their interrelationships at the spine and hip. J Bone Miner Metab 20:358–366

50. Deng HW, Mahaney MC, Williams JT, Li J, Conway T, Davies KM, Li JL, Deng H, Recker RR (2002) Relevance of the genes for bone mass variation to susceptibility to osteoporotic fractures and its implications to gene search for complex human diseases. Genet Epidemiol 22:12–25

51. Deng HW, Chen WM, Recker S, Stegman MR, Li JL, Davies KM, Zhou Y, Deng H, Heaney R, Recker RR (2000) Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. J Bone Miner Res 15:1243–1252

52. Neath AA, Cavanaugh JE (2006) A Bayesian approach to the multiple comparisons problem. J Data Sci 4:131–146

53. Cheverud JM, Routman EJ, Irschick DJ (1997) Pleiotropic effects of individual gene loci on mandibular morphology. Evolution 51:2006–2016

54. Garcia-Cortes LA, Cabrillo C, Moreno C, Varona L (2001) Hypothesis testing for the genetic background of quantitative traits. Genet Sel Evol 33:3–16

55. Raftery A, Lewis S (1992) How many iterations in the Gibbs sampler? In: Bernardo J, Berger J, Dawid A, Smith A (eds) Bayesian statistics. Oxford University Press, Oxford, United Kingdom

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

57. Yi N, George V, Allison DB (2003) Stochastic search variable selection for identifying multiple quantitative trait loci. Genetics 164:1129–1138

58. Xu S (2003) Estimating polygenic effects using markers of the entire genome. Genetics 163:789–801

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

60. Mao Y, Xu S (2005) A Monte Carlo algorithm for computing the IBD matrices using incomplete marker information. Heredity 94:305–315

61. Wang T, Fernando RL, Grossman M (1998) Genetic evaluation

by best linear unbiased prediction using marker and trait information in a multibreed population. Genetics 148:507–515

62. Meuwissen TH, Karlsen A, Lien S, Olsaker I, Goddard ME (2002) Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics 161:373–379

63. Meuwissen THE, Goddard ME (2001) Prediction of identity-by-descent probabilities from marker-haplotypes. Genet Sel Evol 33:605-634

64. Allison DB, Thiel B, St Jean P, Elston RC, Infante MC, Schork NJ (1998) Multiple phenotype modeling in gene-mapping studies of quantitative traits: power advantages. Am J Hum Genet 63:1190–1201

65. Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longmans Green Harlow, Essex, United Kingdom