A tutorial on Reversible Jump MCMC with a view toward applications in QTL-mapping

Rasmus Waagepetersen* Daniel Sorensen†

Abstract

A tutorial derivation of the reversible jump Markov chain Monte Carlo (MCMC) algorithm is given. Various examples illustrate how reversible jump MCMC is a general framework for Metropolis-Hastings algorithms where the proposal and the target distribution may have densities on spaces of varying dimension. It is finally discussed how reversible jump MCMC can be applied in genetics to compute the posterior distribution of the number, locations, effects, and genotypes of putative quantitative trait loci.

Keywords: Markov chain Monte Carlo; Metropolis-Hastings algorithm; reversible jump; quantitative trait loci mapping.

1 Introduction

Markov chain Monte Carlo (MCMC) has become a very important computational tool in Bayesian statistics, since it allows for Monte Carlo approximation of complex posterior distributions where analytical or numerical integration techniques are not applicable. MCMC has also been applied, to a lesser degree, for Monte Carlo approximation of likelihoods; see Geyer and Thompson (1992). The well-known Metropolis (Metropolis et al., 1953) and Metropolis-Hastings (Hastings, 1970) MCMC algorithms were introduced for simulation of target distributions on a space of fixed dimension as e.g. n dimensional Euclidean space. The reversible jump MCMC algorithm introduced in Green (1995), allows simulation from target distributions on spaces of varying dimension. Reversible jump MCMC can also be considered as a general framework for Metropolis-Hastings algorithms, where the proposal distribution and the target distribution have densities on spaces of different dimension. Both simultaneous and single-site updating Metropolis-Hastings algorithms are in fact special cases of reversible jump MCMC.

*Biometry Research Unit, Department of Agricultural Systems, Danish Institute of Agricultural Sciences, PB 50, DK-8830 Tjele, Denmark. E-mail: Rasmus.Waagepetersen@agrsci.dk

†Section on Biometrical Genetics, Department of Animal Breeding and Genetics, Danish Institute of Agricultural Sciences, PB 50, DK-8830 Tjele, Denmark. E-mail: sorensen@inet.uni2.dk

The purpose of this work is to present a simple, self-contained derivation of the reversible jump MCMC algorithm, using a minimum of measure theory. Our intention is to make the topic accessible to biometrical geneticists, quantitative biologists in general, and applied statisticians, following the tutorial style in Casella and George (1992) and Chib and Greenberg (1995). The paper is organized as follows. In section 2, a few examples motivate and illustrate possible applications of reversible jump MCMC. In section 3, a short review of reversible Markov chain Monte Carlo algorithms is given. Section 4 focuses on reversible jump MCMC, and finally, section 5 contains an example of how reversible jump MCMC might be used in genetics for Bayesian mapping of quantitative trait loci (QTL).

2 Statistical models with varying dimensions

Different models with parameter spaces of varying dimensionality may be postulated for the analysis of a particular data set. A natural approach to take model uncertainty into account in the Bayesian framework, is to assign a prior distribution over the collection of competing models. The posterior distribution over the collection of models and the unknown model parameters cannot be handled within the usual Metropolis-Hastings framework and calls for application of the reversible jump MCMC methodology. We give here a few examples of the huge number of situations where reversible jump computation has been applied:

Example 2.1 *Mixtures with unknown number of components.*

Suppose that $f_1(\cdot;\theta_1),\ldots,f_k(\cdot;\theta_k)$ are densities and that data are assumed to be generated from the mixture density $f=\sum_{i=1}^{k}\delta_i f_i$. Apart from the mixture weigths δ_i and the parameters θ_i, $i=1,\ldots,k$, one may also be interested in inference concerning the number k of densities appearing in the mixture. Richardson and Green (1997) apply reversible jump MCMC to carry out Bayesian inference for mixture models with unknown number of mixture components.

Example 2.2 *Nonparametric Bayesian smoothing.*

Suppose that f is a smooth function on an interval $[a,b]\subseteq\mathbb{R}$ and that data $Y_i=f(x_i)+\epsilon_i$, $i=1,\ldots,n$, are observed, where the x_i's are fixed locations in $[a,b]$ and the ϵ_i's are *iid* zero-mean normal variables. A nonparametric estimate of f may be obtained by fitting a step-function to the observations. In a Bayesian approach one may introduce priors for the unknown number N of locations for the jumps of the step-function, the locations $X_1,\ldots,X_N$ of the jumps, and the heights $H_1,\ldots,H_N$ of each step, and reversible jump MCMC can be used to sample the posterior distribution over the space of step-functions with varying number of steps. The posterior mean typically has a smooth appearance and can be used as an estimate of f. The variable dimension Bayesian approach to curve fitting was first proposed in Arjas and Gasbarra (1994). A related approach is considered in Denison et al. (1998).

Example 2.3 *Linear regression with varying number of covariates.*

One may here envisage a number of competing regression models with different number of covariates. In the QTL (Quantitative Trait Loci)-mapping example at the end of this paper, the data are assumed to follow a linear normal model where the dimension of the design matrix depends on the unknown number of QTLs.

Example 2.4 *Finite point processes.*

A finite point process on a bounded subset $D \subseteq \mathbb{R}^d$, is a finite set $\{X_1, \ldots, X_N\}$ of points X_i randomly distributed in D. If the number N of points is random one may need to apply a reversible jump MCMC algorithm in order to simulate the point process. The MCMC algorithm introduced in Geyer and Møller (1994) for simulation of finite point processes is a special case of the reversible jump algorithm. In many situations where reversible jump MCMC is applied to Bayesian inference, the priors and posteriors may on the other hand naturally be regarded as finite point processes. In example 2.2, for instance, the random set of jump locations is a finite point process on $[a, b]$, and the set $\{(X_1, H_1), \ldots, (X_N, H_N)\}$ of jump locations together with the associated step heights constitute a so-called marked point process with marks given by the step heights. A similar remark holds for the QTL example in section 5.

3 Reversible Markov chain Monte Carlo

3.1 Prerequisites

3.1.1 Notation

Suppose that $Z = (Z_1, \ldots . Z_d)$ is a real random vector of dimension $d \geq 1$. We shall say that Z has the density f on $\mathbb{R}^d$ if the probability that Z belongs to a subset A of $\mathbb{R}^d$ is given by the d-fold integral

$$P(Z \in A) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} 1((z_1, \ldots, z_d) \in A) f(z_1, \ldots, z_d) dz_1 \cdots dz_d$$

where the indicator function $1((z_1, \ldots, z_d) \in A)$ is one if $(z_1, \ldots, z_d)$ is in A, and zero otherwise. If e.g. $A = [a_1, b_1] \times [a_2, b_2]$ is a two-dimensional rectangle then the integral is:

$$\int \int 1(x \in [a_1, b_1], y \in [a_2, b_2]) f(x, y) dx dy = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dx dy.$$

Integrals will usually be written in abbreviated form:

$$P(Z \in A) = \int 1(z \in A) f(z) dz = \int_A f(z) dz$$

whenever the dimension of the integral is clear from the context.

3.1.2 Conditional distributions

We shall further need the notion of a conditional distribution. Suppose that X and Y are stochastic vectors in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively, and that X has the density f on $\mathbb{R}^p$. In general a function $0 \leq \mathsf{P}(\cdot,\cdot) \leq 1$ which takes as its first argument elements in $\mathbb{R}^p$ and as its second argument subsets of $\mathbb{R}^q$ is said to be a conditional distribution of Y given X when the identity

$$P(X \in A, Y \in B) = \int_A \mathsf{P}(x, B) f(x) dx \tag{1}$$

holds for all subsets $A \subseteq \mathbb{R}^p$ and $B \subseteq \mathbb{R}^q$.

If (X, Y) has a joint density h on $\mathbb{R}^{p+q}$ then the conditional distribution of Y given $X = x$ is given by the familiar expression

$$\mathsf{P}(x, B) = P(Y \in B | X = x) = \int_B h(y|x) dy$$

where $B \subseteq \mathbb{R}^q$ and $h(y|x) = h(x, y)/f(x)$ is the conditional density on $\mathbb{R}^q$ for Y. The identity (1) is then just the well-known identity $P(X \in A, Y \in B) = \int_A \int_B h(y|x) f(x) dy dx$.

Later on, in section 3.3 and section 4, we shall consider situations where, conditionally on $X = x$, Y is given by $Y = g(x, U)$ where g is a deterministic mapping and U is a stochastic variable with a density $q(x, \cdot)$ on $\mathbb{R}^{q'}$ $(q' < q)$. Then $\mathsf{P}(x, B)$ is given by

$$\mathsf{P}(x, B) = \int 1(g(x, u) \in B) q(x, u) du.$$

That is, conditionally on $X = x$, Y does not have a density on $\mathbb{R}^q$ but on a space of lower dimension. In this case we cannot write the joint probability $P(X \in A, Y \in B)$ in terms of a joint density on $\mathbb{R}^{p+q}$ but need to refer to the equation (1).

3.2 A short review of MCMC

In the following it is assumed that Z has density π on $\mathbb{R}^d$, where π has a complex form, so that expectations with respect to π cannot be evaluated analytically or by using standard techniques for numerical integration. In particular, π may only be known up to an unknown normalizing constant. Direct simulation of π may be difficult, but usually it is quite easy to construct a Markov chain whose *invariant* (or *stationary*) distribution is given by π. The Markov chain $X_1, X_2, X_3, \ldots$ is specified in terms of the distribution for the initial state X_1 and the transition kernel $\mathsf{P}(\cdot,\cdot)$ which specifies the conditional distribution of X_{i+1} given the previous state X_i. That is, if the value of the current state is $X_i = x$, then the probability that X_{i+1} is in a set $A \subseteq \mathbb{R}^d$ is given by

$$\mathsf{P}(x, A) = P(X_{i+1} \in A | X_i = x).$$

Under the conditions of irreducibility and aperiodicity, the generated Markov chain becomes ergodic and can be used for Monte Carlo estimation of various expectations $E(h(Z))$

with respect to the invariant density π. That is, for almost all initial values of X_1 and any function h on $\mathbb{R}^d$ with finite expectation $E(h(Z))$,

$$E(h(Z)) = \int h(z)\pi(z)dz = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} h(X_i). \tag{2}$$

Thus, $E(h(Z))$ can be approximated by the sample average $\sum_{i=1}^{N} h(X_i)/N$ for some large N. Here, h could for instance be the indicator function of a set $A \subseteq \mathbb{R}^d$, so that $E(h(Z))$ equals the probability $P(Z \in A) = E(1(Z \in A))$. For details on convergence of MCMC estimates we refer the reader to e.g. Chib and Greenberg (1995) or Besag et al. (1995).

The distribution π is invariant for the Markov chain $(X_i)_{i\geq 1}$ if the transition kernel $\mathsf{P}(\cdot,\cdot)$ of the Markov chain preserves π, i.e. if $X_i \sim \pi$ implies $X_{i+1} \sim \pi$, or, in terms of $\mathsf{P}(\cdot,\cdot)$ and π, if

$$\int \mathsf{P}(x, B)\pi(x)dx = \int_B \pi(x)dx \tag{3}$$

for any subset $B \subseteq \mathbb{R}^d$. Imposing the stronger condition of *reversibility* with respect to π, is sufficient to guarantee that π is invariant for the Markov chain. This condition holds if (X_i, X_{i+1}) has the same distribution as the time-reversed subchain (X_{i+1}, X_i) whenever $X_i \sim \pi$, i.e. if

$$P((X_i, X_{i+1}) \in A \times B) = \int_A \mathsf{P}(x, B)\pi(x)dx = \int_B \mathsf{P}(x, A)\pi(x)dx = P((X_i, X_{i+1}) \in B \times A) \tag{4}$$

for subsets $A, B \subseteq \mathbb{R}^d$. Reversibility (4) implies (3) by taking $A = \mathbb{R}^d$ (since $\mathsf{P}(x, \mathbb{R}^d) = 1$).

To verify that a non-reversible chain has the correct invariant distribution is in general a difficult task, since this involves integration in (3) with respect to π, and difficulties with this was the reason for using MCMC in the first place. However, the Metropolis-Hastings algorithm, or more generally reversible jump MCMC, offers a practical recipe for constructing a reversible Markov chain with a given invariant distribution.

3.3 Simultaneous and single-site updating Metropolis-Hastings algorithms

One motivation for reversible jump MCMC is the requirement for simulation of posterior distributions on spaces of varying dimension. Reversible jump may be seen also as a general framework for Metropolis-Hastings algorithms with degenerate proposal distributions, i.e. proposal distributions which do not have a density on the space where the target density is defined. To illustrate this, a comparison is made between the well-known simultaneous and single-site updating Metropolis-Hastings algorithms.

In the following, X_n denotes the current n'th state of a Metropolis-Hastings chain $X_1, X_2, \ldots$ for some $n \geq 1$, and Y_{n+1} denotes the proposal for the next state of the chain.

The proposal Y_{n+1} is either accepted with a certain acceptance probability which depends on the values of X_n and Y_{n+1}, or it is rejected, in which case the next state X_{n+1} of the Metropolis-Hastings chain is equal to X_n. We return to the acceptance probability in section 4.4.

For a joint updating Metropolis-Hastings algorithm in equilibrium, the random vector (X_n, Y_{n+1}) consisting of the current Markov chain state and the proposal, has joint density g on $\mathbb{R}^{2d}$, given by

$$g(x, y) = q(x, y)\pi(x)$$

where π is the d-dimensional target density, and $q(x, \cdot)$ is the d-dimensional proposal density of Y_{n+1}, given that X has the value $x \in \mathbb{R}^d$. The probability that Y_{n+1} belongs in a set $A \subseteq \mathbb{R}^d$ given that $X_n = x$, is thus given by the following integral on $\mathbb{R}^d$:

$$\mathsf{Q}(x, A) = P(Y_{n+1} \in A | X_n = x) = \int 1(y \in A) q(x, y) dy. \tag{5}$$

The situation is different for a single-site updating Metropolis-Hastings algorithm for which only one component of the current state X_n is updated at a time. Suppose that Y_{n+1} is obtained by updating the i'th component of $X_n = x$. Then, Y_{n+1} equals x except at the i'th component, where x_i is replaced by a random variable U generated from a one-dimensional proposal density $q_i(x, \cdot)$. Since

$$Y_{n+1} \in A \Leftrightarrow (x_1, \ldots, x_{i-1}, U, x_{i+1}, \ldots, x_d) \in A,$$

the probability that Y_{n+1} belongs in $A \subseteq \mathbb{R}^d$, given that $X_n = x$, is given by the one-dimensional integral:

$$\mathsf{Q}(x, A) = P(Y_{n+1} \in A | X_n = x) = \int 1\left((x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_d) \in A\right) q_i\left(x, u\right) du. \tag{6}$$

Due to the different forms of expressions (5) and (6) for the proposal probabilities $\mathsf{Q}(x, A)$, simultaneous and single-site algorithms are usually treated separately in the literature. Both the joint and the single-site updating Metropolis-Hastings algorithms are, however, special cases of the reversible jump algorithm, which we turn to in the next section.

4 Reversible jump MCMC

In Green (1995), the reversible jump algorithm is defined in a very general setting, and it is not immediately clear how to apply the methodology in a specific situation. Here we give a detailed derivation of the algorithm in a framework which covers many situations of interest as e.g. example 2.1 and example 2.2 and where the notation is still manageable. The approach should hopefully enable the reader to handle also situations not covered by the chosen framework (see e.g. section 5.1).

4.1 The invariant distribution on continuous spaces of varying dimension

In the following, π is the joint probability distribution of (M, Z), where $M \in \{1, 2, 3, \ldots, I\}$ is a "model indicator" and Z is a real stochastic vector of possibly varying dimension (I either represents a finite integer or ∞). The vector Z takes values in the set C defined as the union $C = \cup_{m=1}^{I} C_m$ of spaces $C_m = \mathbb{R}^{n_m}$, $n_m \geq 1$. Given $M = m$, Z can only take values in C_m, so that π is specified by $p_m = P(M = m)$ and densities $f(\cdot|M = m)$ on C_m, $m = 1, 2, \ldots$ for the distribution of Z given $M = m$. Thus, for $A_m \subseteq C_m$:

$$P(M = m, Z \in A_m) = P(M = m)\, P(Z \in A_m | M = m) = p_m \int_{A_m} f(x|M = m)\, dx.$$

The density $f(\cdot|M = m)$ is denoted f_m hereinafter.

If one contemplates a number of competing models, where the number of parameters may vary across models, then p_m may represent the posterior probability of model m, and given $M = m$, f_m is the posterior density of the n_m-dimensional vector Z of parameters associated with model m. In this case $p_m f_m(z) = c^{-1} \tilde{p}_m h(z|m) g(y|m, z)$ where $\tilde{p}_m$ is the prior probability of model m, $h(z|m)$ is the prior density of z given $M = m$, $g(y|m, z)$ is the likelihood of the data y given $(M, Z) = (m, z)$, and c is the overall (typically unknown) normalizing constant $c = \sum_{m=1}^{I} \tilde{p}_m \int_{C_m} g(y|m, z) h(z|m) dz$.

4.2 Specification of the algorithm

We now construct a reversible Markov chain $(X_i)_{i \geq 1}$ with invariant distribution π in a manner similar to the Metropolis-Hastings algorithm. Each state X_i contains two components, i.e. $X_i = (M_i, Z_i)$, where M_i is the model indicator, and Z_i is a stochastic vector in C_{M_i}. Suppose that (m, z) is the value of the current state X_n of the Markov chain. A proposal $Y_{n+1} = (Y_{n+1}^{\text{ind}}, Y_{n+1}^{\text{par}})$, is generated as follows for the new state X_{n+1} of the Markov chain, where the superscripts *ind* and *par* are labels for the proposals of the model indicator M_{n+1} and of the vector Z_{n+1}, respectively. With probability $p_{mm'}$ ($\sum_{m'=1}^{I} p_{mm'} = 1$), the proposal Y_{n+1}^{ind} for the new model indicator is set equal to m', and given $Y_{n+1}^{\text{ind}} = m'$, the proposal Y_{n+1}^{par} is generated in $C_{m'}$. It may be useful to generate Y_{n+1}^{par} by applying a deterministic mapping to the previous value z and to a random component U. A general formulation for this mechanism is to express Y_{n+1}^{par} as $Y_{n+1}^{\text{par}} = g_{1mm'}(z, U)$, where U is a random vector on $\mathbb{R}^{n_{mm'}}$, $n_{mm'} \geq 1$, which has density $q_{mm'}(z, \cdot)$ on $\mathbb{R}^{n_{mm'}}$, and $g_{1mm'} : \mathbb{R}^{n_m + n_{mm'}} \longrightarrow \mathbb{R}^{n_{m'}}$ is a deterministic mapping. The proposal Y_{n+1} is finally accepted with an acceptance probability $a_{mm'}(z, Y_{n+1}^{\text{par}})$ to be derived in section 4.4.

When considering a move from a state (m, z) to $(m', z') = (m', g_{1mm'}(z, u))$, and the reverse move from (m', z') to $(m, z) = (m, g_{1m'm}(z', u'))$, the vectors of Markov chain states and proposal random variables (z, u) and (z', u'), must be of equal dimension. That is, the crucial *dimension matching condition*

$$n_m + n_{mm'} = n_{m'} + n_{m'm} \tag{7}$$

needs to be fulfilled. This condition ensures that $f_m(z)q_{mm'}(z,u)$ and $f_{m'}(z')q_{m'm}(z',u')$ are joint densities on spaces of equal dimension.

It will further be assumed that there exist functions $g_{2mm'} : \mathbb{R}^{n_m+n_{mm'}} \to \mathbb{R}^{n_{m'm}}$ and $g_{2m'm} : \mathbb{R}^{n_{m'}+n_{m'm}} \to \mathbb{R}^{n_{mm'}}$, such that the mapping $g_{mm'}$ given by

$$(z', u') = g_{mm'}(z, u) = (g_{1mm'}(z, u), g_{2mm'}(z, u)) \tag{8}$$

is one-to-one with

$$(z, u) = g^{-1}_{mm'}(z', u') = g_{m'm}(z', u') = (g_{1m'm}(z', u'), g_{2m'm}(z', u',)) \tag{9}$$

and that $g_{mm'}$ is differentiable.

Example 4.1 *Simultaneous and single-site updating random walk Metropolis.*

Let f be a density on $\mathbb{R}^d$, and let $I = 1$, $C_1 = \mathbb{R}^d$, $p_1 = 1$, $f_1 = f$, and $p_{11} = 1$ in the reversible jump framework. The simultaneous updating random walk Metropolis algorithm for simulation of f is then obtained with $g_{11}(z, U) = g(z, U) = (z + U, -U)$ where U is generated from a density $q(\cdot)$ on $\mathbb{R}^d$. The transition kernel for updating of the i'th component in a single-site updating random walk Metropolis algorithm is obtained with $g_{11}(z, U) = (z_1, \ldots, z_{i-1}, z_i + U, z_{i+1}, \ldots, z_d, -U)$ where U is generated from a density on $\mathbb{R}$.

Example 4.2 *Jumps between two models of equal dimension.*

A simple example with two models of equal dimension is as follows. Let $I = 2$, $C_1 = \mathbb{R}_+ \times \mathbb{R}_+$ and $C_2 = \mathbb{R} \times \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the positive real numbers. Further, let $f_1(\alpha, \beta|y) \propto g_1(y|\alpha, \beta)h_1(\alpha, \beta)$, where g_1 is a gamma density and h_1 is a prior for the gamma density parameters α and β. Also, let $f_2(\mu, \sigma^2|y) \propto g_2(y|\mu, \sigma^2)h_2(\mu, \sigma^2)$, where g_2 is a log normal density and h_2 is a prior for the log normal density parameters μ and σ^2. Suppose that the current state of the Markov chain is $(1, \alpha, \beta)$, and that a move is to be made to the log normal model, i.e. to a state $(2, \mu, \sigma^2)$. One way to propose values for the parameters μ and σ^2 might be to equate the first and second order moments under the current gamma model and the proposed log normal model, and subsequently add/multiply some noise. More precisely, solve $\exp(\tilde{\mu} + \tilde{\sigma}^2/2) = \alpha\beta$ and $\exp(2\tilde{\mu} + 2\tilde{\sigma}^2) = \beta^2\alpha(\alpha + 1)$ with respect to $\tilde{\mu}$ and $\tilde{\sigma}^2$, and let the proposals be $\mu = \tilde{\mu} + U$ and $\sigma^2 = \tilde{\sigma}^2 V$ where U and V are generated from q_{12}. One thereby obtains

$$g_{12}(\alpha, \beta, U, V) = \left(\log(\alpha\beta/\sqrt{1 + 1/\alpha}) + U, \log(1 + 1/\alpha)V, U, V\right)$$

and

$$g_{21}(\mu, \sigma^2, U', V') = \left(1/(\exp(\sigma^2/V') - 1), \exp(\mu - U' + \sigma^2/(2V'))(\exp(\sigma^2/V') - 1), U', V'\right).$$

4.3 Reversibility

Assuming $X_n = (M_n, Z_n) \sim \pi$, the condition of reversibility is

$$P(M_n = m, Z_n \in A_m, M_{n+1} = m', Z_{n+1} \in B_{m'}) = P(M_n = m', Z_n \in B_{m'}, M_{n+1} = m, Z_{n+1} \in A_m) \quad (10)$$

for all $m, m' \in \{1, \ldots, I\}$, and all subsets A_m and $B_{m'}$ of C_m and $C_{m'}$, respectively. In analogy with (1), the left hand side of (10) is

$$P\left(M_n = m, Z_n \in A_m, M_{n+1} = m', Z_{n+1} \in B_{m'}\right) = p_m \int_{A_m} f_m(z)\, P\left(M_{n+1} = m', Z_{n+1} \in B_{m'} | X_n = (m, z)\right) dz.$$

Let

$$\mathsf{Q}^a_{mm'}(z, B_{m'}) = P\left(Y^{\text{ind}}_{n+1} = m', Y^{\text{par}}_{n+1} \in B_{m'} \text{ and } Y_{n+1} \text{ is accepted} | X_n = (m, z)\right)$$

be the joint probability of generating a proposal Y_{n+1} with $Y^{\text{ind}}_{n+1} = m'$ and Y^{par}_{n+1} in $B_{m'}$ and accepting the proposal, given that the current state of the Markov chain is $X_n = (m, z)$. Let further

$$s_m(z) = P\left(Y_{n+1} \text{ is rejected} | X_n = (m, z)\right) = \sum_{m'=1}^{I} p_{mm'} \int q_{mm'}(z, u) \left[1 - a_{mm'}\left(z, g_{1mm'}(z, u)\right)\right] du$$

be the probability of rejecting the proposal. Then

$$P\left(M_{n+1} = m', Z_{n+1} \in B_{m'} | X_n = (m, z)\right) = \mathsf{Q}^a_{mm'}(z, B_{m'}) + s_m(z)\, 1\left(m = m', z \in B_{m'}\right).$$

The left hand side of (10) therefore equals

$$p_m \int_{A_m} f_m(z)\, \mathsf{Q}^a_{mm'}(z, B_{m'})\, dz + p_m \int_{A_m} f_m(z)\, s_m(z)\, 1\left(m = m', z \in B_{m'}\right) dz = \int_{A_m} p_m f_m(z)\, \mathsf{Q}^a_{mm'}(z, B_{m'})\, dz + \int p_m f_m(z)\, s_m(z)\, 1\left(m = m', z \in A_m \cap B_{m'}\right) dz \quad (11)$$

where

$$\int_{A_m} p_m f_m(z)\, \mathsf{Q}^a_{mm'}(z, B_{m'})\, dz = P\left(M_n = m, Z_n \in A_m, Y^{\text{ind}}_{n+1} = m', Y^{\text{par}}_{n+1} \in B_{m'} \text{ and } Y_{n+1} \text{ is accepted}\right).$$

By symmetry the right hand side of (10) equals

$$\int_{B_{m'}} p_{m'} f_{m'}(z') \, \mathsf{Q}^a_{m'm}(z', A_m) \, dz' + \int p_{m'} f_{m'}(z') \, s_{m'}(z') \, 1(m = m', z' \in B_{m'} \cap A_m) \, dz'. \tag{12}$$

The second terms in (11) and (12) are equal both in the case when $m \neq m'$ (in which case they are zero, because the indicator function takes the value zero), and when $m = m'$ (in which case the move is within the same model, and both expressions are identical). Therefore a sufficient condition for (10) to hold is, for all m and m':

$$\int_{A_m} p_m f_m(z) \, \mathsf{Q}^a_{mm'}(z, B_{m'}) \, dz = \int_{B_{m'}} p_{m'} f_{m'}(z') \, \mathsf{Q}^a_{m'm}(z', A_m) \, dz'. \tag{13}$$

4.4 Derivation of the acceptance probability

Equation (13) is now written more explicitly. Since (a) Y_{n+1} is generated in $C_{m'}$ with probability $p_{mm'}$, (b) $Y_{n+1} \in B_{m'} \Leftrightarrow g_{1mm'}(z, U) \in B_{m'}$, (c) Y_{n+1} is accepted with probability $a_{mm'}(z, g_{1mm'}(z, U))$, and (d) $U \sim q_{mm'}(z, .)$, it follows that

$$\mathsf{Q}^a_{mm'}(z, B_{m'}) = p_{mm'} \int 1(g_{1mm'}(z, u) \in B_{m'}) \, a_{mm'}(z, g_{1mm'}(z, u)) \, q_{mm'}(z, u) \, du \tag{14}$$

and the left hand side of (13) is thus

$$\int_{A_m} p_m f_m(z) \, \mathsf{Q}^a_{mm'}(z, B_{m'}) \, dz =$$
$$\int\int 1(z \in A_m, g_{1mm'}(z, u) \in B_{m'}) \, p_m f_m(z) \, p_{mm'} a_{mm'}(z, g_{1mm'}(z, u)) \, q_{mm'}(z, u) \, dz du. \tag{15}$$

Similarly, the right hand side of (13) is

$$\int_{B_{m'}} p_{m'} f_{m'}(z') \, \mathsf{Q}^a_{m'm}(z', A_m) \, dz' =$$
$$\int\int 1(z' \in B_{m'}, g_{1m'm}(z', u') \in A_m) \, p_{m'} f_{m'}(z') \, p_{m'm} a_{m'm}(z', g_{1m'm}(z', u')) \, q_{m'm}(z', u') \, dz' du'. \tag{16}$$

Due to the dimension matching assumption (7) and the relationships (8) and (9) involving $g_{mm'}$ which is assumed to be differentiable, a straightforward change of variable in (16) yields:

$$\int\int 1(g_{1mm'}(z, u) \in B_{m'}, z \in A_m) \, p_{m'} f_{m'}(g_{1mm'}(z, u)) \, p_{m'm}$$
$$a_{m'm}(g_{1mm'}(z, u), z) \, q_{m'm}(g_{1mm'}(z, u), g_{2mm'}(z, u)) \, |g'_{mm'}(z, u)| dz du \tag{17}$$

where $dz'du' = |g'_{mm'}(z,u)|dzdu$ and

$$g'_{mm'}(z,u) = \frac{\partial g_{mm'}(z,u)}{\partial z \partial u} = \begin{bmatrix} \frac{\partial g_{1mm'}(z,u)}{\partial z} & \frac{\partial g_{2mm'}(z,u)}{\partial z} \\ \frac{\partial g_{1mm'}(z,u)}{\partial u} & \frac{\partial g_{2mm'}(z,u)}{\partial u} \end{bmatrix}.$$

Examination of (15) and of (17) shows that the reversibility condition (13) is satisfied if

$$p_m f_m(z)\, p_{mm'} q_{mm'}(z,u)\, a_{mm'}(z, g_{1mm'}(z,u)) =$$
$$p_{m'} f_{m'}(g_{1mm'}(z,u))\, p_{m'm} q_{m'm}(g_{1mm'}(z,u), g_{2mm'}(z,u))\, a_{m'm}(g_{1mm'}(z,u), z) \left| \frac{\partial g_{mm'}(z,u)}{\partial z \partial u} \right|. \tag{18}$$

Choosing the acceptance probability as large as possible subject to the detailed balance condition (18), as suggested by Peskun (1973), yields:

$$a_{mm'}(z,z') = \min\left(1, \frac{p_{m'} f_{m'}(z')\, p_{m'm} q_{m'm}(z',u')}{p_m f_m(z)\, p_{mm'} q_{mm'}(z,u)} \left| \frac{\partial g_{mm'}(z,u)}{\partial z \partial u} \right| \right) \tag{19}$$

whenever $p_m f_m(z)\, p_{mm'} q_{mm'}(z,u) > 0$ and where $(z',u') = g_{mm'}(z,u)$. In practice, $p_m f_m(z)\, p_{mm'} q_{mm'}(z,u) = 0$ only happens if the Markov chain is initialized in a state (m,z) for which $p_m f_m(z) = 0$.

Remark 4.1 *The Jacobian.*

Note that the Jacobian in (19) appears due to the deterministic transformation used in the proposal mechanism, and the change of variable used when equating (15) and (16). The Jacobian is therefore not really an inherent component of dimension changing MCMC; in many situations with varying dimension it actually equals one (see e.g. section 5). In the fixed dimension example 4.2 the Jacobian on the other hand equals $|\partial g_{12}/\partial\alpha\partial\beta\partial u\partial v| = v(\beta\alpha(\alpha+1))^{-1}$ which differs from one.

Remark 4.2 *Deterministic proposals.*

Sometimes it may be useful to apply deterministic proposals for a move from C_m to $C_{m'}$, i.e. to let $Y_{n+1} = g_{1mm'}(z)$, and still use a stochastic proposal for the move in the other direction. In this case $n_{mm'}$ equals zero and the dimension-matching condition becomes $n_m = n_{m'} + n_{m'm}$. A function $g_{2mm'} : \mathbb{R}^{n_m} \to \mathbb{R}^{n_{m'm}}$ is thus required, such that the inverse of $g_{mm'} = (g_{1mm'}, g_{2mm'})$ is given by $g_{m'm} = g_{1m'm}$. In other words, equations (8) and (9) become

$$(z',u') = g_{mm'}(z) = (g_{1mm'}(z), g_{2mm'}(z)) \text{ and } z = g^{-1}_{mm'}(z',u') = g_{1m'm}(z',u').$$

Equation (14) becomes

$$\mathrm{Q}^a_{mm'}(z, B_{m'}) = p_{mm'} 1(g_{1mm'}(z) \in B_{m'})\, a_{mm'}(z, g_{1mm'}(z)), \tag{20}$$

and calculations similar to those leading to (19) yield the acceptance probability

$$a_{mm'}(z,z') = \min\left(1, \frac{p_{m'} f_{m'}(z')\, p_{m'm} q_{m'm}(z',u')}{p_m f_m(z)\, p_{mm'}} \left| \frac{\partial g_{mm'}(z)}{\partial z} \right| \right).$$

5 Bayesian mapping of QTL

Geneticists are often interested in locating regions in the chromosome contributing to phenotypic variation of a quantitative trait. These chromosomal regions are known as quantitative trait loci (QTL), and their location and effects on the quantitative trait can be investigated using genetic markers. The latter are chromosomal regions of known location in the chromosome, which show detectable DNA variation among individuals in the population, and which do not have a physiological causal association to the trait under study. By studying the joint pattern of inheritance of the markers and the trait, inferences can be made about the number, location and effects of the QTLs affecting the trait. A Bayesian approach combined with reversible jump MCMC is well suited for QTL studies, and applications can be found in Satagopan and Yandell (1996), Heath (1997), Uimari and Hoeschele (1997), Stephens and Fisch (1998), and Sillanpää and Arjas (1998). Here we consider a simple experimental design in order to illustrate the use of reversible jump MCMC in QTL-mapping.

Assume that $K > 1$ linked flanking markers and phenotypic observations $y_i, i = 1, \ldots, n_{\text{obs}}$ are available from a backcross population of n_{obs} individuals. For simplicity, it is assumed that full marker information is available for all individuals. These individuals were generated by mating the offspring from the cross between two completely inbred lines, back to parents taken from one of the two inbred lines. In this backcross design, at any one locus or gene location, the genotypic distribution comprises only two possible types. The marker genotype information for the K marker loci for each of the n_{obs} individuals is denoted $\mathcal{M} = \{\mathcal{M}_{ij}\}_{i=1,\ldots,n_{\text{obs}}, j=1,\ldots,K}$, and $D = \{D_l\}_{l=1}^{K}$ denotes the known positions of the K markers. Suppose that m QTLs are present at locations $\lambda_1, \ldots, \lambda_m$, where $D_1 < \lambda_i < D_K$, and let $o_{ij} \in \{0, 1\}$ denote the genotype of the i'th individual for the QTL at location λ_j. More than one QTL may be present in the region defined by two flanking markers. Let Q be the matrix whose j'th column is $o_j = (o_{1j}, \ldots, o_{n_{\text{obs}}j})^{\text{T}}$, the vector of QTL-genotypes at the location λ_j for all n_{obs} individuals. The data $y = (y_1, \ldots, y_{n_{\text{obs}}})^{\text{T}}$ is assumed to be a realized value from Y, where

$$Y|m, \lambda_1, \ldots, \lambda_m, \beta_1, \ldots, \beta_m, Q, \mu, \sigma^2 \sim N\left((\mu, \ldots, \mu)^{\text{T}} + Q(\beta_1, \ldots, \beta_m)^{\text{T}}, \sigma^2 I\right). \qquad (21)$$

In (21), $\beta_1, \ldots, \beta_m$ are real parameters describing the effects of each QTL, μ is an intercept, and $\sigma^2 > 0$ is a residual variance which may include a polygenic contribution from other loci affecting the trait. For ease of presentation it is assumed that μ and σ^2 are known.

The Bayesian model requires the specification of a prior distribution for the model parameters. Given $M = m$, the locations $\lambda_1, \ldots, \lambda_m$ are assumed to be independent and uniformly distributed in the interval $D =]D_1, D_K[$. For the unknown number M of QTLs a Poisson distribution is assigned, with mean $\alpha(D_K - D_1)$, so that $\alpha > 0$ is the a priori expected number of QTLs in an interval of unit length. Given the QTL-locations $\lambda = (\lambda_1, \ldots, \lambda_m)$, the probability of a genotype configuration $o = (o_{ij})_{i=1,\ldots,n_{\text{obs}},\, j=1,\ldots,m} \in \{0, 1\}^{n_{\text{obs}}m}$ is $p(o|\lambda, \mathcal{M}, m)$. Here we shall not be concerned with the explicit expression for $p(o|\lambda, \mathcal{M}, m)$, but refer the interested reader to e.g. Sillanpää and Arjas (1998). The

QTL effect parameters $\beta_1, \ldots, \beta_m$ have a priori independent, zero-mean, normal distributions with variance σ_0^2. Given $M = m$, the parameters (λ, o) and $\beta = (\beta_1, \ldots, \beta_m)$ are independent, a priori.

The joint posterior is of the form:

$$\begin{aligned} f(\lambda, \beta, o, m|y) &\propto f(y|\lambda, \beta, o, m)\, f(\lambda, \beta, o, m) \\ &= f(y|\beta, o, m; \mu, \sigma^2) p(o|\lambda, m, \mathcal{M}) f(\lambda|m; D) h(\beta|m; \sigma_0^2) p(m; \alpha). \quad (22) \end{aligned}$$

5.1 Reversible jump MCMC for QTL-mapping

In this section it is shown how a reversible jump algorithm for simulation of the joint posterior (22) may be constructed. The posterior does not fall into the framework of section 4 due to the discrete QTL genotype parameters and we therefore briefly explain how the reversible jump algorithm is derived for the current model, in analogy with the derivation in section 4. Attention is restricted to dimension changing moves which either increase or decrease the number of QTLs. Suppose that the current state of the Markov chain has m QTLs. With probability $p_{m,m-1} = 1/2$ it is proposed to decrease the number of QTLs by one, and with probability $p_{m,m+1} = 1/2$ it is proposed to increase the number of QTLs by one (the chain remains at the current state if $m = 0$ and it is proposed to decrease the number of QTLs).

To economise on notation in the following, the state of the Markov chain is represented as (m, z), where $z = (z_1, \ldots, z_m)$ is a vector of QTL-configurations $z_j = (\lambda_j, \beta_j, o_j)$, $j = 1, \ldots, m$. Thus, each QTL-configuration z_j consists of the QTL location together with the associated QTL effect and genotypes. A QTL configuration then belongs in the space $C_{\text{conf}} = D \times \mathbb{R} \times \{0, 1\}^{n_{\text{obs}}}$, and the vector z of m QTL configurations belongs in the space $C_m = C_{\text{conf}}^m$.

5.1.1 Removal of a QTL

Suppose that there are $m \geq 1$ QTL configurations in the current state $X_n = (m, z)$ of the Markov chain. A move which reduces the number of QTLs may be accomplished by simply deterministically removing the last QTL-configuration in z, that is, letting the proposal for the QTL-parameters be $Y_{n+1}^{\text{par}} = (z_1, \ldots, z_{m-1})$. Suppose that A_m is a subset of C_{conf}^m and that B_{m-1} is a subset of C_{conf}^{m-1}. Since the proposal is generated deterministically, the probability $P(Y_{n+1}^{\text{ind}} = m - 1, Y_{n+1}^{\text{par}} \in B_{m-1}$ and Y_{n+1} accepted$|X_n = (m, z))$ is in analogy with (20),

$$\mathsf{Q}_{m,m-1}^a(z, B_{m-1}) = \frac{1}{2} 1((z_1, \ldots, z_{m-1}) \in B_{m-1}) a_{m,m-1}(z, (z_1, \ldots, z_{m-1})) \quad (23)$$

and the joint probability $P(M_n = m, Z_n \in A_m, Y_{n+1}^{\text{ind}} = m-1, Y_{n+1}^{\text{par}} \in B_{m-1}$ and Y_{n+1} accepted) is

$$\int_{D^m} \int_{\mathbb{R}^m} \sum_{o \in \{0,1\}^{n_{\text{obs}} m}} f(z, m|y) 1(z \in A_m) \mathsf{Q}_{m,m-1}^a(z, B_{m-1}) d\lambda d\beta =$$
$$\frac{1}{2} \int_{D^m} \int_{\mathbb{R}^m} \sum_{o \in \{0,1\}^{n_{\text{obs}} m}} f(z, m|y) 1(z \in A_m, (z_1, \ldots, z_{m-1}) \in B_{m-1})$$
$$a_{m,m-1}(z, (z_1, \ldots, z_{m-1})) d\lambda d\beta \quad (24)$$

where $d\lambda d\beta$ is short for $d\lambda_1 \cdots d\lambda_m d\beta_1 \cdots d\beta_m$ and $z = ((\lambda_1, \beta_1, o_1), \ldots, (\lambda_m, \beta_m, o_m))$.

5.1.2 Addition of a QTL

Suppose now, conversely, that the current state is $(m-1, z') = (m-1, z_1, \ldots, z_{m-1})$ and that the number of QTLs is to be increased by one. Then, the position λ_m of the new QTL is sampled uniformly between D_1 and D_K. The genotypes o_m for the new QTL are sampled from the conditional probability $p(o_m|\lambda', o', \lambda_m, \mathcal{M}, m)$ of the QTL genotypes at location λ_m, given the existing QTL locations λ', the existing genotypes o', and the new location λ_m. The new regression parameter β_m is finally sampled from $N(0, \tau^2)$, $\tau^2 > 0$. The proposal for the QTL parameters is then $z = (z_1, \ldots, z_{m-1}, (\lambda_m, \beta_m, o_m))$. In analogy with (14), the probability $P(Y_{n+1}^{\text{ind}} = m, Y_{n+1}^{\text{par}} \in A_m$ and Y_{n+1} accepted$|X_n = (m-1, z'))$ is

$$\mathsf{Q}_{m-1,m}^a(z', A_m) =$$
$$\frac{1}{2} \int_D \int_{\mathbb{R}} \sum_{o_m \in \{0,1\}^{n_{\text{obs}}}} 1((z', (\lambda_m, \beta_m, o_m)) \in A_m) a_{m-1,m}(z', (z', (\lambda_m, \beta_m, o_m)))$$
$$q_{m-1,m}(z', (\lambda_m, \beta_m, o_m)) \, d\lambda_m d\beta_m \quad (25)$$

where

$$q_{m-1,m}(z', (\lambda_m, \beta_m, o_m)) = f(\beta_m; \tau^2) \left(\frac{1}{D_K - D_1} \right) p(o_m|\lambda', o', \lambda_m, \mathcal{M}, m)$$

and analogously to (15), $P(M_n = m-1, Z_n \in B_{m-1}, Y_{n+1}^{\text{ind}} = m, Y_{n+1}^{\text{par}} \in A_m$ and Y_{n+1} accepted) is

$$\int_{D^{m-1}} \int_{\mathbb{R}^{m-1}} \sum_{o' \in \{0,1\}^{n_{\text{obs}}(m-1)}} f(z', m-1|y) 1(z' \in B_{m-1}) \mathsf{Q}_{m-1,m}^a(z', A_m) d\lambda' d\beta' \quad (26)$$

where $d\lambda' d\beta'$ is short for $d\lambda_1 \cdots d\lambda_{m-1} d\beta_1 \cdot \cdots \cdot d\beta_{m-1}$ and $z' = ((\lambda_1, \beta_1, o_1), \ldots,$ $(\lambda_{m-1}, \beta_{m-1}, o_{m-1}))$. Substituting (25) in (26) yields:

$$\frac{1}{2} \int_{D^m} \int_{\mathbb{R}^m} \sum_{o' \in \{0,1\}^{n_{\text{obs}} m}} f(z', m-1|y) \frac{f(\beta_m; \tau^2)}{D_K - D_1} p(o_m|\lambda', o', \lambda_m, \mathcal{M}, m)$$
$$1(z' \in B_{m-1}, (z', (\lambda_m, \beta_m, o_m)) \in A_m) a_{m-1,m}(z', (z', (\lambda_m, \beta_m, o_m))) d\lambda' d\beta' d\lambda_m d\beta_m. \quad (27)$$

5.1.3 The acceptance probability

Note that the dimensions of the integrals and sums in (24) and (27) match. As in section 4, the acceptance probability is derived by equating (24) and (27). The situation is here particularly simple, since we do not need a change of variable. To make the connection to section 4 more clear we could as in Remark 4.2 express the removal-proposal as $z' = g_{1m,m-1}(z_1, \ldots, z_m)$ and the addition-proposal as $z = g_{1m-1,m}(z_1, \ldots, z_{m-1}, (\lambda_m, \beta_m, o_m))$. The functions $g_{m,m-1}$ and $g_{m-1,m}$ are then the identity mappings and the Jacobian arising from a formal change of the continuous QTL location and effect variables equals one. In analogy with Remark 4.2, the acceptance probability for the move from (m, z) to $(m-1, z') = (m-1, z_1, \ldots, z_{m-1})$ can readily be obtained:

$$a_{m,m-1}(z, z') = \min\{1, \frac{f(z', m-1|y) f(\beta_m; \tau^2) p(o_m|\lambda', o', \lambda_m, \mathcal{M}, m)}{f(z, m|y)(D_K - D_1)}\} \qquad (28)$$

where $\lambda' = (\lambda_1, \ldots, \lambda_{m-1})$ and $o' = (o_1, \ldots, o_{m-1})$.

5.1.4 Discussion

For ease of exposition the last configuration in the list of QTL-configurations is deterministically chosen for a proposed removal. One could alternatively as in Sillanpää and Arjas (1998) pick the QTL-configuration to be removed at random or perhaps introduce a new type of move which "shuffles" the order of the QTL-configurations. Such a "shuffle" move does not change the likelihood for the data and would therefore always be accepted. For the same reason it would be natural to identify all vectors $(z_1, \ldots, z_m)$, $(z_2, z_1, \ldots, z_m) \ldots$ which differ only by a permutation with the (unordered) set $\{z_1, \ldots, z_m\}$ which could be considered as a realization of a marked point process on D. The acceptance probability (28) is then actually a special case of the acceptance probability derived in Geyer and Møller (1994) for Metropolis-Hastings simulation of finite point processes.

The new QTL effect β_m in section 5.1.2 is generated in a way similar to the proposal used in Sillanpää and Arjas (1998). A more elaborate method for generating β_m is used in Satagopan and Yandell (1996).

Acknowledgements We wish to thank Elja Arjas, Dario Gasbarra, and Mikko Sillanpää for many helpful comments and suggestions.

References

ARJAS, E. and GASBARRA, D. (1994) Nonparametric Bayesian inference from right censored survival data, using the Gibbs sampler. *Statistica Sinica* **4,** 505-524.

BESAG, J. E., GREEN, P. J., HIGDON, D., and MENGERSEN, K. (1995) Bayesian computation and stochastic systems. *Statist. Sci.* **10,** 3-66.

CASELLA, G. and GEORGE, E. (1992) Explaining the Gibbs sampler. *The American Statistician* **46,** 167-174.

CHIB, S. and GREENBERG, E. (1995) Understanding the Metropolis-Hastings algorithm. *American Statistician* **49,** 329-335.

DENISON, D. G. T., MALLICK, B. K. and SMITH, A. F. M. (1998) Automatic Bayesian curve fitting. *J. Roy. Statist. Soc. B* **60,** 333-350.

GEYER, C. J. and MØLLER, J. (1994) Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Statist.* **21,** 359-373.

GEYER, C. J. and THOMPSON, E. A. (1992) Constrained Monte Carlo maximum likelihood for dependent data (with discussion). *J. Roy. Statist. Soc. B* **54,** 657-699.

GREEN, P. J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82,** 711-732.

HASTINGS, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57,** 97-109.

HEATH, S. C. (1997) Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am. J. Hum. Genet.* **61**, 748-760.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., and TELLER, E. (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **21,** 1087-1092.

PESKUN, P. H. (1973) Optimum Monte-Carlo sampling using Markov-chains. *Biometrika* **60,** 607-612.

RICHARDSON, S. and GREEN, P. J. (1997) On Bayesian analysis of mixtures with an unknown number of components (with discussion). *J. Roy. Statist. Soc. B* **59,** 731-792.

SATAGOPAN, J. M. and YANDELL, B. S. (1996) Estimating the number of quantitative trait loci via Bayesian model determination. Special Contributed Paper Session on: Genetic Analysis of Quantitative Traits and Complex Diseases; Biometric Section, Joint Statistical Meetings, Chicago, IL., USA. (available at ftp://ftp.stat.wisc.edu/pub/yandell/revjump.html/).

SATAGOPAN, J. M., YANDELL, B. S., NEWTON M. A., and OSBORN, T. C. (1996) A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. *Genetics* **144**, 805-816.

SILLANPÄÄ, M. J. and ARJAS, E. (1998) Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* **148**, 1373-1388.

STEPHENS, D. A. and FISCH, R. D. (1998) Bayesian analysis of quantitative trait locus data using reversible jump Markov chain Monte Carlo. *Biometrics* **54,** 1334-1347.

UIMARI, P. and HOESCHELE, I. (1997) Mapping linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. *Genetics* **146**, 735-743.