Genet. Sel. Evol. 30 (1998) 103–125

Original article

Bayesian estimation of dispersion parameters with a reduced animal model including polygenic and QTL effects

Marco C.A.M. Bink [a*], Richard L. Quaas [b], Johan A.M. Van Arendonk [a]

[a] Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen Agricultural University, PO Box 338, 6700 AH Wageningen, the Netherlands
[b] Department of Animal Science Cornell University, Ithaca, NY 14853, USA

(Received 21 April 1997; accepted 29 December 1997)

Abstract – In animal breeding, Markov chain Monte Carlo algorithms are increasingly used to draw statistical inferences about marginal posterior distributions of parameters in genetic models. The Gibbs sampling algorithm is most commonly used and requires full conditional densities to be of a standard form. In this study, we describe a Bayesian method for the statistical mapping of quantitative trait loci (QTL), where the application of a reduced animal model leads to non-standard densities for dispersion parameters. The Metropolis Hastings algorithm is used to obtain samples from these non-standard densities. The flexibility of the Metropolis Hastings algorithm also allows us change the parameterization of the genetic model. Alternatively to the usual variance components, we use one variance component (= residual) and two ratios of variance components, i.e. heritability and proportion of genetic variance due to the QTL, to parameterize the genetic model. Prior knowledge on ratios can more easily be implemented, partly by absence of scale effects. Three sets of simulated data are used to study performance of the reduced animal model, parameterization of the genetic model, and testing the presence of the QTL at a fixed position. © Inra/Elsevier, Paris

reduced animal model / dispersion parameters / Markov chain Monte Carlo / quantitative trait loci

Résumé – Estimation Bayésienne des paramètres de dispersion dans un modèle animal réduit comprenant un effet polygénique et l'effet d'un QTL. En génétique animale, les algorithmes de Monte-Carlo par chaînes de Markov sont utilisés de plus en plus souvent pour en inférer aux distributions marginales a posteriori des paramètres

* Correspondence and reprints
E-mail: marco.bink@alg.vf.wau.nl

du modèle génétique. L'algorithme d'échantillonnage de Gibbs est utilisé largement et demande la connaissance des densités conditionnelles, dans une forme standard. Dans cette étude, on décrit une méthode Bayésienne pour la cartographie statistique d'un locus à effet quantitatif (QTL), où l'application d'un modèle animal réduit conduit à des densités de paramètres de dispersion, qui n'ont pas de forme standard. On utilise l'algorithme de Metropolis-Hastings pour l'échantillonnage de ces densités non standard. La souplesse de l'algorithme de Metropolis-Hastings permet également de changer la paramétrisation du modèle génétique : au lieu des composantes de variances habituelles, on peut utiliser une composante de variance (résiduelle) et deux rapports de composantes de variance : l'héritabilité et la proportion de la variance génétique dûe au QTL. Il est plus facile de spécifier l'information a priori sur des proportions, en partie parce qu'elle ne dépend pas de l'échelle. Trois fichiers de données simulées sont utilisés pour étudier la performance du modèle animal réduit, par rapport au modèle animal strict, l'effet de paramétrisation du modèle génétique et la qualité du test de la présence d'un QTL à une position donnée.

modèle animal réduit / paramètres de dispersion / méthode de Monte-Carlo par chaînes de Markov / locus quantitatif

1. INTRODUCTION

The wide availability of high-speed computing and the advent of methods based on Monte Carlo simulation, particularly those using Markov chain algorithms, have opened powerful pathways to tackle complicated tasks in (Bayesian) statistics [9, 10]. Markov chain Monte Carlo (MCMC) methods provide means for obtaining marginal distributions from a complex non-standard joint density of all unknown parameters (which is not feasible analytically). There are a variety of techniques for implementation [9] of which Gibbs sampling [11] is most commonly used in animal breeding. The applications include univariate models, threshold models, multi-trait analysis, segregation analysis and QTL mapping [15, 17, 29, 31, 33].

Because Gibbs sampling requires direct sampling from full conditional distributions, data augmentation [22] is often used so that 'standard' sampling densities are obtained. Often, however, this is at the expense of a substantial increase in number of parameters to be sampled. For example, the full conditional density for a genetic variance component becomes standard (inverted gamma distribution) when a genetic effect is sampled for each animal in the pedigree, as in a (full) animal model (FAM). The dimensionality increases even more rapidly when the FAM is applied to the analysis of granddaughter designs [34] in QTL mapping experiments, i.e. marker genotypes on granddaughters are not known and need to be sampled as well. In addition, absence of marker data hampers accurate estimation of genetic effects within granddaughters, which form the majority in a granddaughter design. This might lead to very slow mixing properties of the dispersion parameters (see also Sorensen et al. [21]).

The reduced animal model (RAM, Quaas and Pollak, [19]) is equivalent to the FAM, but can greatly reduce the dimensionality of a problem by eliminating effects of animals with no descendants. With a RAM, however, full conditional densities for dispersion parameters are not standard. Intuitively, RAM, used to eliminate genetic effects and concentrate information, is the antithesis of data augmentation, used to arrive at simple standard densities. For the Metropolis-Hastings (MH) algorithm

[14, 18], however, a standard density is not required, in fact, the sampling density needs to be known only up to proportionality. Another alternative for the FAM is the application of a sire model which implies that only sires are evaluated based on progeny records. With a sire model, the genetic merit of the dam of progeny is not accounted for and only the phenotypic information on offspring is used. The RAM offers the opportunity to include maternal relationships, offspring with known marker genotypes and information on grandoffspring. As a result the RAM is better suited for the analysis of data with a complex pedigree structure.

The flexibility of the MH algorithm also allows for a greater choice of the parameterization (variance components or ratios thereof) of the genetic model. If Gibbs sampling is to be employed, the parameterization is often dictated by mathematical tractability to obtain the simple sampling density. The MH algorithm readily admits much flexibility in modelling prior belief regarding dispersion parameters, which is an advantageous property in Bayesian analysis [16].

In this paper, we present MCMC algorithms that allow Bayesian linkage analysis with a RAM. We study two alternative parameterizations of the genetic model and use a test statistic to postulate presence of a QTL at a fixed position relative to an informative marker bracket. Three sets of simulation data using a typical granddaughter design are used.

2. METHOD

2.1. Genetic model

The additive genetic variance (σ_a^2) underlying a quantitative trait is assumed to be due to two independent random effects, due to a putative QTL and residual independent polygenes. The QTL effects ($\mathbf{v}$) are assumed to have a $N(\mathbf{0}, \mathbf{G}\sigma_v^2)$ prior distribution where $\mathbf{G}$ is the gametic relationship matrix [2, 8], and σ_v^2 is the variance due to a single allelic effect at the QTL. Matrix $\mathbf{G}$ depends upon one unknown parameter, the map position of the QTL relative to the (known) positions of bracketing (informative) markers. Here we consider the location of the QTL to be known. The polygenic effects ($\mathbf{u}$) have a $N(\mathbf{0}, \mathbf{A}\sigma_u^2)$ prior distribution, where $\mathbf{A}$ is the numerator relationship matrix. The genetic model underlying the phenotype of an animal is

$$y_i = x_i\mathbf{b} + u_i + v_i^1 + v_i^2 + e_i$$

where $\mathbf{b}$ is the vector with fixed effects, v_i^1 and v_i^2 are the two (allelic) QTL effects for animal i, and $e_i \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$. (QTL effects within individual are assigned according to marker alleles, as proposed by Wang et al. [32]). The sum of the three genetic effects is the animal's breeding value (a). In addition to genetic effects, location parameters comprise fixed effects that are, a priori, assumed to follow the proper uniform distribution: $f(\mathbf{b}) \sim U[\mathbf{b}_{\min}, \mathbf{b}_{\max}]$, where $\mathbf{b}_{\min}$ and $\mathbf{b}_{\max}$ are the minimum and maximum values for elements in $\mathbf{b}$.

2.2. Reduced animal model (RAM)

The RAM is used to reduce the number of location parameters that need to be sampled. The RAM eliminates the need to sample genetic effects of animals

with neither descendants nor marker genotypes, i.e. ungenotyped non-parents. The phenotypic information on these animals can easily be absorbed into their parents without loss of information. Absorption of non-parents that have marker genotypes becomes more complex when position of QTL is unknown; it is therefore better to include them explicitly in the analysis. In the remainder of the paper, it is assumed that marker genotypes on non-parents are not available. The genetic effects of non-parents can be expressed as linear functions of the parental genetic effects by the following equations [4],

$$\mathbf{u}_{\text{non-parents}} = \mathbf{P}_{\text{parents}}\mathbf{u}_{\text{parents}} + \varphi_{\text{non-parents}} \tag{1}$$

and

$$\mathbf{v}_{\text{non-parents}} = \mathbf{Q}_{\text{parents}}\mathbf{v}_{\text{parents}} + \phi_{\text{non-parents}} \tag{2}$$

where each row in $\mathbf{P}$ contains at most two non-zero elements (= 0.5), and each row in $\mathbf{Q}$ has at most four non-zero elements [32], the terms $\varphi_{\text{non-parents}}$ and $\phi_{\text{non-parents}}$ pertain to remaining genetic variance due to Mendelian segregation of alleles. In a granddaughter design, the $\mathbf{P}$ and $\mathbf{Q}$ for granddaughters, not having marker genotypes observed nor augmented, have similar structures,

$$\mathbf{Q} = \mathbf{P} \otimes \frac{1}{2}\mathbf{J}_{2\times 2} \tag{3}$$

where $\otimes$ denotes the Kronecker product, and $\mathbf{J}$ is a unity matrix [20]. This equality does not hold if marker genotypes are augmented, since phenotypes contain information that can alter the marker genotype probabilities for ungenotyped non-parents [2].

The phenotypes for a quantitative trait can now be expressed as,

$$y_i = x_i\mathbf{b} + \mathbf{P}_i\mathbf{u} + \mathbf{Q}_i\mathbf{v} + \varepsilon_i \tag{4}$$

for row vectors $\mathbf{P}_i$ and $\mathbf{Q}_i$ (possibly null), and

$$\sigma^2_{\varepsilon_i} = \sigma^2_{\text{e}} + \omega_i(\sigma^2_{\text{u}} + 2\sigma^2_{\text{v}}) \tag{5}$$

where ω_i reflects the amount of total additive genetic variance that is present in $\sigma^2_{\varepsilon_i}$. Based on the pedigree, four categories of animals are distinguished in the RAM (*table I*). The vectors $\mathbf{P}_i$ and $\mathbf{Q}_i$ contain partial regression coefficients. For parents, the only non-zero coefficients pertain to the individual's own genetic effects (ones); for non-parents, the individual's parents' genetic effects (halves). Note that $\mathbf{P}_i$ and $\mathbf{Q}_i$ are null for a non-parent with unknown parents, and that non-parents' phenotypes in this category contribute to the estimation of fixed effects and phenotypic (residual) variance only.

2.3. Parameterization

Let θ denote the set of location parameters ($\mathbf{b}$, $\mathbf{u}$ and $\mathbf{v}$) and dispersion parameters.

Table I. Categories of animals in a reduced animal model and values for ω_i for each category.

Category		No. of parents known	ω_I^a
1	non-parent	0	1
2	non-parent	1	3/4
3	non-parent	2	1/2
4	parent	–[b]	0

[a] Without inbreeding; [b] not relevant.

We consider the following two parameterizations for the dispersion parameters

$$\theta_{VC}: \quad \mathbf{b},\ \mathbf{u},\ \mathbf{v},\ \sigma_e^2,\ \sigma_u^2 \text{ and } \sigma_v^2$$

$$\theta_{RT}: \quad \mathbf{b},\ \mathbf{u},\ \mathbf{v},\ \sigma_e^2,\ h^2 \text{ and } \gamma$$

where

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2} \quad \text{or} \quad \frac{\sigma_u^2 + 2\sigma_v^2}{\sigma_e^2 + \sigma_u^2 + 2\sigma_v^2}, \quad \text{with} \quad 0 \leqslant h^2 \leqslant 1 \tag{6}$$

and

$$\gamma = \frac{2\sigma_v^2}{\sigma_a^2} \quad \text{or} \quad \frac{2\sigma_v^2}{\sigma_u^2 + 2\sigma_v^2}, \quad \text{with} \quad 0 \leqslant \gamma \leqslant 1 \tag{7}$$

In the first, θ_{VC}, the parameters are the variance components (VC). This is the usual parameterization. A difficulty with this is that it is problematic for an animal breeder to elicit a reasonable prior of the genetic VC. Animal breeders, it seems to us, are much more likely to have, and be able to state, prior opinions about such things as heritabilities. Consequently, in θ_{RT}, parameter h^2 is the heritability of a trait, and parameter γ is the proportion of additive genetic variance due to the putative QTL. This parameterization allows more flexible modelling of prior knowledge because h^2 and γ do not depend on scale. Theobald et al. [23] used a variance ratio, σ_u^2/σ_e^2, parameterization but noted that the animal breeder may prefer to think in terms of heritability. We prefer the part-whole ratios h^2 and γ. The components σ_u^2 and σ_v^2 can be expressed in terms of σ_e^2, h^2 and γ

$$\sigma_u^2 = (1-\gamma)\frac{h^2}{(1-h^2)}\,\sigma_e^2 \tag{8}$$

and

$$\sigma_v^2 = (0.5 \times \gamma)\frac{h^2}{(1-h^2)}\,\sigma_e^2 \tag{9}$$

2.4. Priors

We now present the prior knowledge on dispersion parameters, priors for location parameters having been given earlier. In earlier studies, two different priors are often

used to describe uncertainty on VC. The inverted gamma (IG) distribution, or its special case the inverted chi-square distribution, is common because it is often the conjugate prior for the VC if the FAM (or sire model) is applied. Hence, the full conditional distribution for VC will then be a posterior updating of a standard prior [9]. This simplifies Gibbs sampling. We will use the IG as the prior for σ_x^2 – though with a RAM it is not conjugate,

$$f(\sigma_x^2|\alpha_x, \beta_x) \propto (\sigma_x^2)^{-\alpha_x - 1} \exp\left\{-\frac{1}{\beta_x}\frac{1}{\sigma_x^2}\right\} \tag{10}$$

where x = e, u, or v. The rhs of (10) constitutes the kernel of the distribution. The mean (μ) of an IG(α, β) is $((\alpha - 1)\beta)^{-1}$, and the variance equals $\left((\alpha - 1)^2(\alpha - 2)\beta^2\right)^{-1}$. Van Tassell et al. [29] suggest setting $\alpha = 2.000001$ and $\beta \approx (\mu)^{-1}$ for an 'almost flat' prior with a mean corresponding to prior expectation (μ). The IG distributions for three different prior expectations are given in *figure 1*. When the prior expectation is close to zero ($\mu = 5.0$), the distribution is more peaked and has less variance because mass accumulates near zero. When the prior expectation is relatively high ($\mu = 60$), the probability of σ_x^2 being equal to zero is very small, which might be undesirable and/or unrealistic for σ_v^2. An alternative prior distribution for σ_x^2 is

$$f(\sigma_x^2) \propto \begin{cases} k_x & 0 \leqslant \sigma_x^2 \leqslant \sigma_{x,\,\max}^2 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

which is a proper prior for σ_x^2 with a uniform density over a pre-defined large, finite interval, for example from zero to 200 (*figure 1*). These prior distributions for VC are used mainly to represent prior uncertainty [21, 30, 31].

Corresponding to (10) (11) there is an equivalent prior distribution for $\lambda(\gamma)$.

However, because neither (10) nor (11) were chosen for any intrinsic 'rightness' we prefer a simpler alternative of using Beta distributions for the ratio parameters λ and γ to represent prior knowledge,

$$f(x|\alpha_x, \beta_x) \propto (x)^{\alpha_x - 1)}(1 - x)^{\beta_x - 1} \tag{12}$$

where $x = h^2$ or γ. When prior distribution parameters α_x and β_x are both set equal to 1, the prior is a uniform density between 0 and 1 (*figure 2*), i.e. flat prior. Alternatively, α_x and β_x can be specified to represent prior expectations for parameters of interest (*figure 2*). For example, one can centre the density for heritability of a yield trait in dairy cattle around the prior expectation (= 0.40), with a relatively flat (Beta (2.5, 3.75)) or peaked (Beta (30.0, 45.0)) distribution when prior certainty is moderate or strong, respectively. Furthermore, prior knowledge on γ, proportion of additive genetic variance due to a putative QTL, can be modelled to give relatively high probabilities of values close to zero, e.g. (Beta (0.9, 2.7)). Another option, suggested by a reviewer, would be to put vague priors on α_x and β_x as in Berger [1].

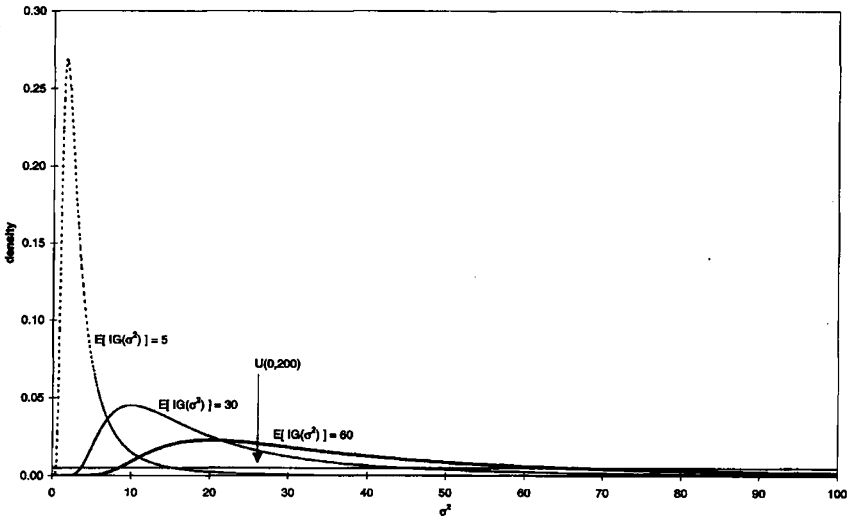

Figure 1. Inverted gamma and uniform densities that are used to represent (lack of) prior knowledge on variance components.

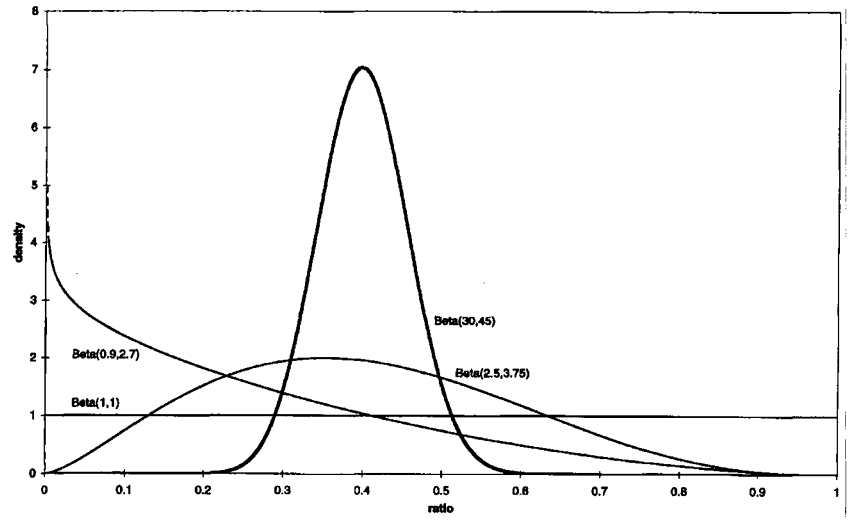

Figure 2. Beta densities that are used to represent (lack of) prior knowledge on (part whole) ratios of variance components.

2.5. Joint posterior density

The joint posterior density of θ is the product of likelihood and prior densities of elements in θ, described above. Let n_i denote the number of observations on

animals of category i (*table I*), the total number of observations being given as N, and let q denote the number animals with offspring, i.e. parents. Then, $2q$ is the number of QTL effects (two allelic effects per animal). With θ_{VC},

$$f(\theta_{\mathrm{VC}}|\mathbf{y},\, \alpha_{\mathrm{e}},\, \beta_{\mathrm{e}},\, \alpha_{\mathrm{u}},\, \beta_{\mathrm{u}},\, \alpha_{\mathrm{v}},\, \beta_{\mathrm{v}}) \;\propto\; f(\theta_{\mathrm{VC}},\, \mathbf{y}|\alpha_{\mathrm{e}},\, \beta_{\mathrm{e}},\, \alpha_{\mathrm{u}},\, \beta_{\mathrm{u}},\, \alpha_{\mathrm{v}},\, \beta_{\mathrm{v}})$$

$$\begin{aligned}
&\propto \prod_{i=1}^{4}\left[\left(\sigma_{\mathrm{e}}^2+\omega_i(\sigma_{\mathrm{u}}^2+2\sigma_{\mathrm{v}}^2)\right)^{-0.5n_i}\times\exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{n_i}e_k^2/(\sigma_{\mathrm{e}}^2+\omega_i(\sigma_{\mathrm{u}}^2+2\sigma_{\mathrm{v}}^2))\right)\right\}\right]\\
&\quad\times(\sigma_{\mathrm{u}}^2)^{-0.5q}\times\exp\left\{-\frac{1}{2}(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u})\times\frac{1}{\sigma_{\mathrm{u}}^2}\right\}\\
&\quad\times(\sigma_{\mathrm{v}}^2)^{-0.5(2q)}\times\exp\left\{-\frac{1}{2}(\mathbf{v}^T\mathbf{G}^{-1}\mathbf{v})\times\frac{1}{\sigma_{\mathrm{v}}^2}\right\}\\
&\quad\times(\sigma_{\mathrm{e}}^2)^{-\alpha_{\mathrm{e}}-1}\exp\left\{\frac{-1}{\beta_{\mathrm{e}}\sigma_{\mathrm{e}}^2}\right\}\times(\sigma_{\mathrm{u}}^2)^{-\alpha_{\mathrm{u}}-1}\exp\left\{\frac{-1}{\beta_{\mathrm{u}}\sigma_{\mathrm{u}}^2}\right\}\times(\sigma_{\mathrm{v}}^2)^{-\alpha_{\mathrm{v}}-1}\exp\left\{\frac{-1}{\beta_{\mathrm{v}}\sigma_{\mathrm{v}}^2}\right\}
\end{aligned}\tag{13}$$

Under θ_{RT}, dispersion parameters, and priors thereof, are different from θ_{VC}; the joint posterior density is

$$f(\theta_{\mathrm{RT}}|\mathbf{y},\, \alpha_{\mathrm{e}},\, \beta_{\mathrm{e}},\, \alpha_{\mathrm{h}^2},\, \beta_{\mathrm{h}^2},\, \alpha_{\gamma},\, \beta_{\gamma}) \;\propto\; f(\theta_{\mathrm{RT}},\, \mathbf{y}|\alpha_{\mathrm{e}},\, \beta_{\mathrm{e}},\, \alpha_{\mathrm{h}^2},\, \beta_{\mathrm{h}^2},\, \alpha_{\gamma},\, \beta_{\gamma})$$

$$\begin{aligned}
&\sim(\sigma_{\mathrm{e}}^2)^{-0.5N}\times\prod_{i=1}^{4}\left[\left(1+\omega_i\frac{h^2}{1-h^2}\right)^{-0.5n_i}\right.\\
&\quad\left.\times\exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{n_i}\mathrm{e}_k^2\Big/\left(\left(1+\omega_i\frac{h^2}{1-h^2}\right)\right)\times\frac{1}{\sigma_{\mathrm{e}}^2}\right\}\right]\\
&\quad\times\left((1-\gamma)\times\left(\frac{h^2}{1-h^2}\right)\times\sigma_{\mathrm{e}}^2\right)^{-0.5q}\\
&\quad\times\exp\left\{-\frac{1}{2}(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u})\times\frac{1}{(1-\gamma)\times\left(\frac{h^2}{1-h^2}\right)\times\sigma_{\mathrm{e}}^2}\right\}\\
&\quad\times\left((0.5\gamma)\times\frac{h^2}{1-h^2}\times\sigma_{\mathrm{e}}^2\right)^{-0.5(2q)}\\
&\quad\times\exp\left\{-\frac{1}{2}(\mathbf{v}^T\mathbf{G}^{-1}\mathbf{v})\times\frac{1}{(0.5\gamma)\times\frac{h^2}{1-h^2}\times\sigma_{\mathrm{e}}^2}\right\}\\
&\quad\times(\sigma_{\mathrm{e}}^2)^{-\alpha_{\mathrm{e}}-1}\exp\left\{\frac{-1}{\beta_{\mathrm{e}}\sigma_{\mathrm{e}}^2}\right\}\times\left(h^2\right)^{\alpha_{h^2}-1}\left(1-h^2\right)\beta_{h^2}^{-1}\times(\gamma)^{\alpha_\gamma-1}(1-\gamma)^{\beta_\gamma-1}
\end{aligned}\tag{14}$$

2.6. Full conditional densities

From the joint posterior densities (13) and (14), the full conditional density for each element in θ can be derived by treating all other elements in θ as constants and selecting the terms involving the parameter of interest. When this leads to the kernel of a standard density, e.g. Normal for location parameters or an IG distribution, e.g. variance components with FAM, Gibbs sampling is applied to draw samples for that element in θ. Otherwise, the full conditional density is non-standard and sampling needs to be done by other techniques. (All full conditional densities are given in the Appendix).

2.7. Sampling non-standard densities by Metropolis-Hastings algorithm

Sampling a non-standard density can be carried out a variety of ways, including various rejection sampling techniques [6, 7, 12, 13], and Metropolis-Hastings sampling within Gibbs sampling [6]. We use the Metropolis-Hastings algorithm (MH). Let $\pi(x)$ denote the target density, the non-standard density of a particular element in θ, and let $q(x, y)$ be the candidate generating density. Then, the probability of move from current value x to candidate value y for θ_i is,

$$\alpha(x,y) = \begin{cases} \min\left[\dfrac{\pi(y)\,q(y,x)}{\pi(x)\,q(x,y)}, 1\right] & \text{if } \pi(x)\,q(x,y) > 0 \\ 1 & \text{otherwise} \end{cases}$$

When y is not accepted, the value for θ_i remains equal to x, at least until the next update for θ_i. Chib and Greenberg [6] described several candidate generating densities for MH. We use the random walk approach in which candidate y is drawn from a distribution centred around the current value x. To ensure that all sampled parameters are within the parameter space the sampling distribution, $q(x, y)$, was $U(B_L, B_U)$ with

$$B_L = \max(0,\, x - t) \quad \text{for } \sigma_e^2,\, \sigma_u^2,\, \sigma_v^2,\, h^2,\, \gamma$$

$$B_U = \begin{cases} x + t & \text{for } \sigma_e^2,\, \sigma_u^2,\, \sigma_v^2 \\ \min(1, x + t) & \text{for } h^2,\, \gamma \end{cases}$$

where t is a positive constant determined empirically for each parameter to give acceptance rates between 25 and 50 % [6, 24]. For each of the non-standard densities, a univariate MH was used. We perform univariate MH iterations (ten times) within a MCMC cycle to enhance mixing in the MCMC chain, as suggested by Uimari et al. [26].

2.8. Comparison to a full animal model (FAM)

From the conditional densities presented, two hybrid MCMC chains can be used to obtain samples of all unknown parameters (θ_{VC} or θ_{RT}) using a RAM. For comparison, the equivalent FAM can be used with similar parameterization (θ_{VC}

and θ_{RT}). The conditional densities for the FAM are a special case of RAM (see *table I*): all animals are in category 4 and $\omega_i = 0$. In case of θ_{VC} the conditional densities for σ^2_v, σ^2_u, and σ^2_v are now recognizable IG distributions and Gibbs sampling can be used to draw samples from these densities directly. In the case of θ_{RT} the conditional densities for h^2 and γ remain non-standard and MH is used to draw samples. *Table II* gives the four constructed MCMC sampling schemes.

Table II. Sampling algorithms for location and dispersion parameters for alternative models (RAM versus FAM) and parameterizations (θ_{VC} versus θ_{RT}).

	RAM		FAM	
	θ_{VC}	θ_{RT}	θ_{VC}	θ_{RT}
β	GS[a]	GS	GS	GS
u	GS	GS	GS	GS
v	GS	GS	GS	GS
σ^2_e	MH[b]	GS	GS	GS
σ^2_u	MH		GS	
σ^2_v	MH		GS	
h^2		MH		MH
γ		MH		MH

[a] GS = Gibbs sampling ; [b] MH = Metropolis-Hastings algorithm.

2.9. Post MCMC analysis

Depending on the dispersion parameterization (θ_{VC} or γ_{RT}), three of five parameters were sampled *(table II)*. In each MCMC cycle, however, the remaining two were computed, using (6) and (7) or (8) and (9), to allow comparison of results of different parameterizations. For parameter X, the auto-correlation of a sequence of samples was calculated as $\frac{1}{m}\sum_{i=1}^{m-1}[(x_i - \hat{\mu}_x)(x_{i+1} - \hat{\mu}_x)]/\hat{s}^2_x$ where m = number of samples, $\hat{\mu}_x$ and $\hat{s}_x$ are posterior mean and standard deviation, respectively. The correlation among samples for parameters x and z, within MCMC cycles, was computed as $\frac{1}{m}\sum_{i=1}^{m}[(x_i - \hat{\mu}_c)(z_i - \hat{\mu}_z)]/[(\hat{s}_x\hat{s}_z]$. For each parameter an effective sample size (ESS) was computed which estimates the number of independent samples with information content equal to that of the dependent samples [21].

The null hypothesis that $\gamma = 0$ – the QTL explains no genetic variance – was tested via an odds ratio $\frac{\text{mode}\{p(\gamma)\}}{p(\gamma = 0)} > 20$ following Janss et al. [17]. They suggest that this criterion, however, may be quite stringent. The 90 % highest posterior density regions (HPD90) [5], were also computed for parameter γ.

3. SIMULATION

In this study, granddaughter designs were generated by Monte Carlo simulation. The unrelated grandsire families each contained 40 sires that were half sibs. The number of families was 20 except in simulation III where designs with 50 families were simulated as well (*table III*). Polygenic and QTL effects for grandsires were sampled from $N(0, \sigma^2_u)$ and $N(0, \sigma^2_v)$, respectively. The polygenic effect for sires was simulated as $U_S = \frac{1}{2}(u_{GS}) + \Phi_i$, where u_{GS} is the grandsire's polygenic effect, and Φ_i, Mendelian sampling, is distributed independently as $N(0, \mathrm{Var}(\Phi_i))$ with $\mathrm{Var}(\Phi_i) = 0.75 \times \sigma^2_u$ (no inbreeding). Each sire inherited one QTL at random from its (grand) sire. The maternally inherited QTL effect for a sire was drawn from $N(0, \sigma^2_v)$. Each sire had 100 daughters with phenotypes observed, that were generated as

$$y \sim N\left\{0.5\, u_{sire} + \rho v^1_{sire} + (1-\rho) v^2_{sire},\quad 0.75\, \sigma^2_u + \sigma^2_v + \sigma^2_e\right\}$$

where ρ is a 0/1 variable. In all simulations the phenotypic variance and the heritability of the trait were 100 and 0.40, respectively. The proportion of genetic variance due to the QTL ($= \gamma$) was by default 0.25, or 0.10 in simulation III (*table III*). Two genetic markers bracketing the QTL position at 10cM (Haldane mapping function) were simulated with five alleles at each marker, with equal frequencies over alleles per marker. For grandsires, the marker genotypes were fully informative, i.e. heterozygous, and the linkage phase between marker alleles is assumed to be known a priori. The uncertainty on linkage phase in sires can be included in θ, but we did not. All possible linkage phases within sires were weighted by their probability of occurrence and one average relationship matrix between grandsires' and sires' QTL effects was used.

Table III. Simulation of granddaughter designs and MCMC chains.

	Simulation I	Simulation II	Simulation III
No. grandsires	20	20	20, 50
Proportion QTL (γ)[a]	0.25	0.25	0.10, 0.25
No. replicates	1	5	25
Purpose	comparison RAM versus FAM	comparison θ_{VC} versus θ_{RT}	hypothesis testing power for detection
MCMC chains			
Length	500 000	250 000	200 000
Thinning factor	250	250	1 000
Stored samples	2 000	1 000	200

[a] Proportion QTL = proportion of additive genetic variance due to the QTL.

4. RESULTS AND DISCUSSION

4.1. Simulation I: comparison RAM versus FAM

For each of the four MCMC algorithms that are given in *table II*, a single MCMC chain was run and 2 000 thinned samples were used for post-MCMC analysis (*table III*). In the case of θ_{VC}, prior distributions for σ_e^2, σ_u^2 and σ_v^2 were 'flat' IGs (*figure 1*) with expected means equal to 60, 30 and 5 (values used for simulation), respectively. In the case of θ_{RT}, the prior for σ_e^2 was again an IG and priors for h^2 and γ were Beta (2.5, 3.75) and Beta (0.9, 2.7), respectively. *Figure 3* presents the mixing properties for parameter σ_v^2 within the chains for the RAM-θ_{VC} and RAM-θ_{VC} alternatives and points to slower mixing when using the FAM. This slow mixing is also indicated by high auto-correlation (≈ 1) among samples for parameters σ_v^2 and γ when the FAM was used (*table IV*). With the same thinning, the auto-correlation among samples in the RAM is $\leqslant 0.70$. The estimates for posterior mean and coefficient of variation, derived from samples of the four chains, are given in *table V*. These estimates are very similar over models (RAM and FAM) and parameterizations (θ_{VC} and θ_{RT}). The coefficients of variation for σ_v^2 and γ are relatively large and indicate that a posteriori knowledge on these parameters remains small, while estimates for σ_e^2 and h^2 are accurate. The magnitude of the sampling correlation among parameters within MCMC cycles was very similar for both models and parameterizations. The samples for σ_v^2 and σ_u^2 showed a moderately high negative correlation (-0.7), while the sampling correlation between h^2 and γ was relatively low and positive (0.2). The correlation among samples for σ_e^2 and h^2 was very high but apparently did not adversely affect the auto-correlation of these parameters. Taking 100 ESS as a minimum [26] the MCMC chain was rather short for statistical inferences for γ in FAM-θ_{RT}. However, running a longer MCMC chain was not practical since the FAM-θ_{VC} MCMC chain needed 68 593 min CPU (47 days) on a HP 9000–735 (125Hz) workstation. This was almost 100 times the 11 h that were needed to run the RAM with similar chain length.

The slow mixing of parameters for a FAM was likely to be due to the lack of marker data on granddaughters. Distinction between polygenic and QTL effects within these animals is hardly possible. Consequently, they provide little information regarding dispersion but because they are so numerous they dominate the distribution from which the next sample for the dispersion parameter is drawn. Heuristically, one first generates $\mathbf{u}$ and $\mathbf{v}$ with variances reflecting current σ^2. Subsequently one samples a new σ^2 from a peaked distribution with a mean near the sample variance of the $\mathbf{u}$ and $\mathbf{v}$. Not surprisingly one gets back a σ^2 very similar to the previous one, as a result of which the chain is slowly mixing.

The data from simulation I were also used to examine the effect of priors on posterior inferences on the proportion of QTL when θ_{RT} was used. Four different priors for γ were used, ranging from a 'flat' (but not a 'non-informative') uniform prior to a density peaked at zero. The latter reflects the prior expectation that the genetic variance due to the QTL is small or equal to zero. *Figure 4* presents both prior and posterior densities. The uniform and the 'peaked-at-zero' prior resulted in the highest (0.20) and lowest posterior mean estimate (0.10), respectively. For this design, the information from the data is not overwhelming the prior knowledge.

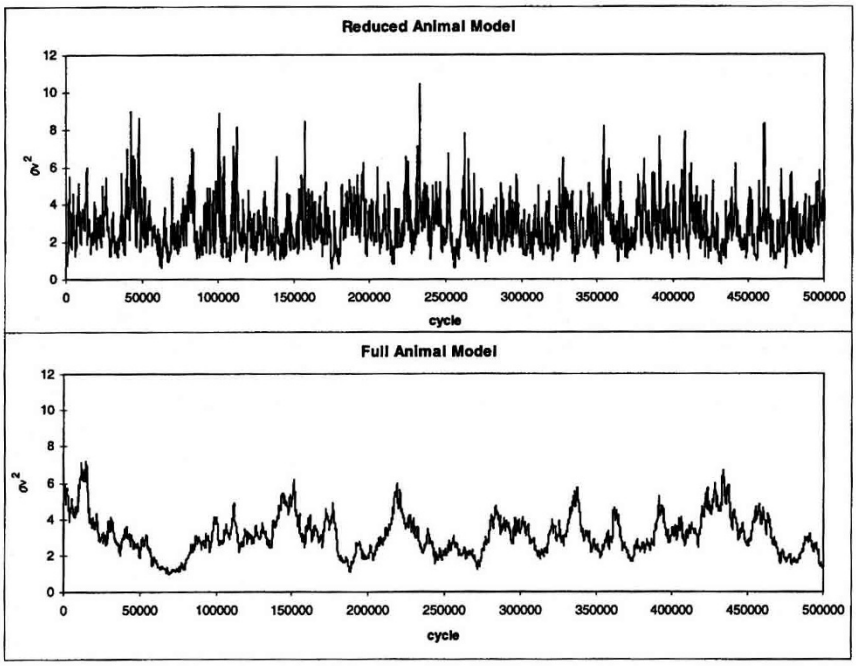

Figure 3. Two-thousand thinned samples for parameter σ_v^2, from MH algorithm (RAM) and from Gibbs sampling (FAM) (simulation I).

Table IV. Sampling correlation and effective sample size for alternative models (RAM versus FAM) and parameterizations (θ_{VC} versus θ_{RT}) from simulation I [a].

		RAM				FAM			
		auto	correlation[b] σ_e^2	σ_u^2	ESS[c]	auto	correlation σ_e^2	σ_u^2	ESS
θ_{VC}	σ_e^2	**0.07**			1 880	**0.29**			1635
	σ_u^2	**0.34**	−0.47		856	**0.61**	−0.57		133
	σ_v^2	**0.60**	−0.29	−0.69	611	**0.97**	−0.18	−0.67	62
θ_{RT}^2	σ_e^2	**0.05**			1 481	**0.57**			654
	h^2	**0.06**	−0.98		1 571	**0.59**	−0.99		604
	γ	**0.71**	−0.19	0.20	350	**0.99**	−0.17	0.17	29

[a] MCMC chains: length 500 000 cycles, initial thinning (k) = 250, samples for analysis (m) = 2 000; [b] auto-correlation = between subsequent samples for the same parameter; otherwise correlation between samples for different parameters within cycle; [c] ESS = effective sample size.

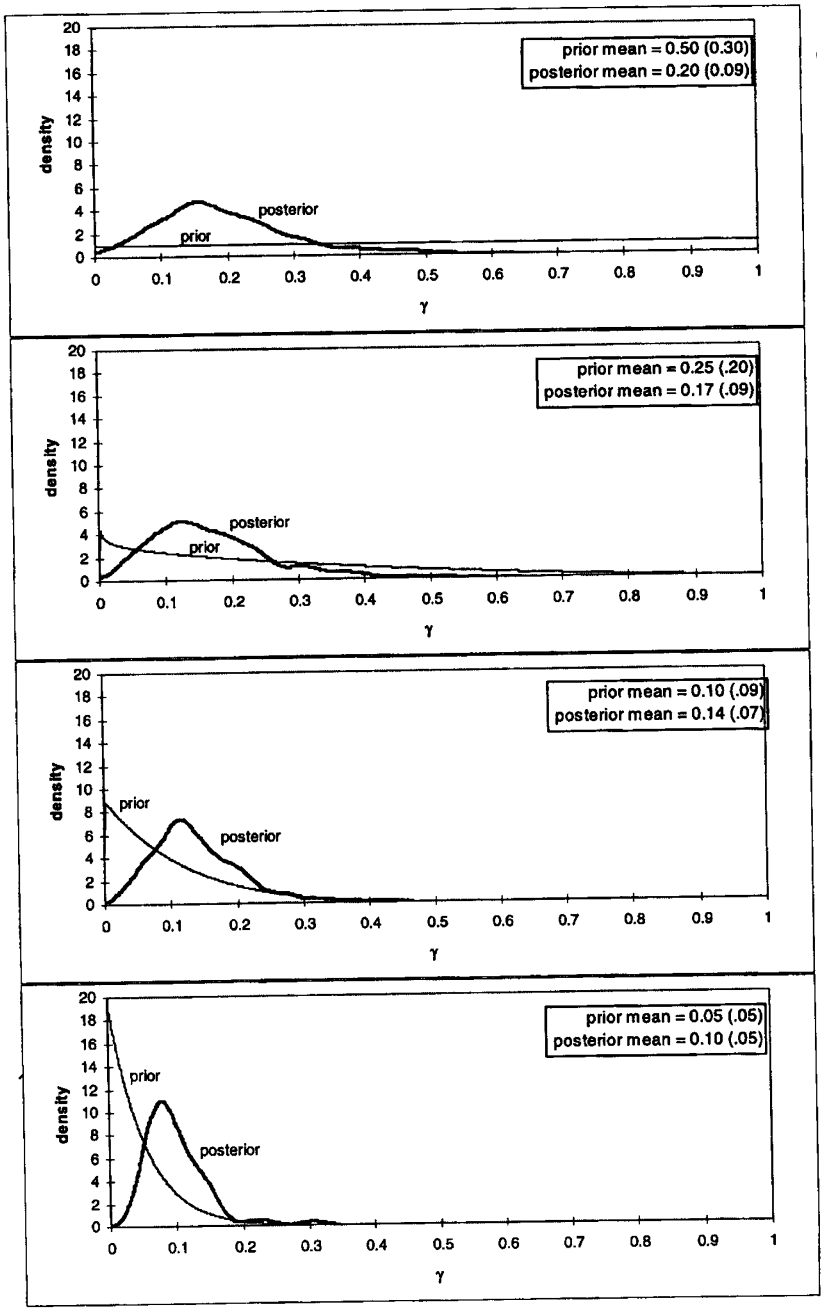

Figure 4. Effect of prior knowledge on posterior densities (RAM – θ_{RT}, simulation I).

Table V. Estimates of posterior mean and standard deviation for dispersion parameters, for alternative models (reduced AM versus full AM) and parameterizations (θ_{VC} versus θ_{RT}) from simulation I[a].

		RAM		FAM	
		mean	CV	mean	CV
θ_{VC}[b]	σ_e^2	62.7	0.03	62.7	0.03
	σ_u^2	30.5	0.09	30.0	0.09
	σ_v^2	2.8	0.44	3.1	0.35
	h^2	0.37	0.05	0.37	0.05
	γ	0.16	0.42	0.17	0.34
θ_{RT}[b]	σ_e^2	62.6	0.03	62.3	0.03
	σ_u^2	30.3	0.11	29.9	0.12
	σ_v^2	3.0	0.53	3.4	0.51
	h^2	0.37	0.05	0.37	0.07
	γ	0.17	0.54	0.18	0.49

[a] MCMC chains: length 500 000 cycles, initial thinning (k) = 250, samples for analysis (m) = 2 000; [b] parameters underlined were actually sampled in that parameterization.

All priors studied, however, showed consistency for the posterior probability of $\gamma = 0$, i.e. the data supported the presence of a QTL at the studied position of the chromosome.

4.2. Simulation II: parameterization of the genetic model

In simulation II, five replicates of data were used to study the effects of alternative parameterizations of the genetic model, for the RAM only. Genetic and population parameters were similar to those in simulation I (*table III*). Based on the results for ESS from the initial MCMC chains (*table IV*), the MCMC chains were run for 250 000 cycles and every 250th was sample used for analysis (m = 1 000). Now, uniform priors for all dispersion parameters were used. The sampling correlations were averaged over the five replicates and are presented in *table VI*. These correlations are consistent with those from the initial MCMC chains (*table IV*); i.e. auto-correlations were highest among samples for σ_v^2 (in θ_{VC}) and γ (in θ_{RT}), i.e. around 0.68. These parameters also had lowest and similar ESS ($\approx$ 230). These results indicate that sampling efficiency is similar for the two studied parameterizations (θ_{VC} and θ_{RT}) of the genetic model and shorter chains may suffice. The posterior mean estimates, averaged over five replicates, for all dispersion parameters were in close agreement with the values used for simulation (not shown).

Table VI. Sampling correlation and effective samples for RAM and alternative parameterizations (θ_{VC} versus θ_{RT}) from simulation II[a].

			RAM correlation[b]		ESS[c]
		auto	σ_e^2	σ_u^2	
	σ_e^2	**0.14**			724
θ_{VC}	σ_u^2	**0.52**	−0.09		284
	σ_v^2	**0.68**	−0.44	−0.84	228
		auto	σ_e^2	h^2	
	σ_e^2	**0.10**			759
θ_{RT}	h^2	**0.11**	−0.99		773
	γ	**0.68**	−0.27	0.28	232

[a] MCMC chains: length 250 000 cycles, initial thinning (k) = 250, samples for analysis (m) = 1 000; [b] auto-correlation = between subsequent samples for the same parameter; otherwise correlation between samples for different parameters within cycle; [c] ESS = effective sample size.

4.3. Simulation III: presence of the QTL

In simulation III, two different designs (20 or 50 grandsire families) were studied in combination with two different sizes of the QTL (explaining either 10 or 25 % of the genetic variance). Two different priors for γ were studied with the θ_{RT} parameterization. For each combination of design and γ, test runs preceding the 25 replicates were used to empirically determine values for t in the MH algorithm, in order to achieve the desired acceptance rates. From the marginal posterior density an odds ratio was computed and the presence of the QTL was accepted only if this ratio exceeded a critical value of 20. Using this test statistic we postulated the power of detecting the QTL for specific designs and using different priors (*table VII*). The small design (20 × 40) has low power of QTL detection, i.e. only 25 %, for a QTL that explains 10 % of the genetic variance. Power increased when either the QTL explained more genetic variance or when a large design (50 × 40) was used. For the large design with a relatively large QTL, power of detection is 100 %, for both priors considered. The use of the 'peaked-at-zero' prior reduced power in the two intermediate cases but increased power in the small design with the small QTL. Estimates for posterior mode, mean and HPD90 were averaged over the 25 replicates and these averages are presented in *figure 5*. When the 'peaked-at-zero' prior was used, point estimates were lower compared to using the uniform prior. This prior also led to shorter – and closer to zero – HPD90 region in all combinations of design and γ but the impact was more noticeable for the small design.

Table VII. Power[a] for detection of QTL for RAM and parameterization θ_{RT} from simulation III[b].

Design[c]	QTL (γ)[d]	prior on γ = Beta (1,1)	prior on γ = Beta (1,19)
20 × 40	0.10	0.24	0.28
	0.25	0.64	0.56
50 × 40	0.10	0.80	0.68
	0.25	1.00	1.00

[a] Power is defined as the acceptance rate for a QTL, for an odds ratio, mode, $\{p(\gamma)\}/p(y = 0)$, exceeds 20. For each 'design QTL' combination, 25 replicates were simulated. [b] MCMC chains: length 200 000 cycles, initial thinning (k) = 1 000, samples for analysis (m) = 200. [c] Design is defined as 20 (50) grandsire families, each family contains 40 sons. [d] QTL (γ) is the proportion of genetic variance due to the QTL.

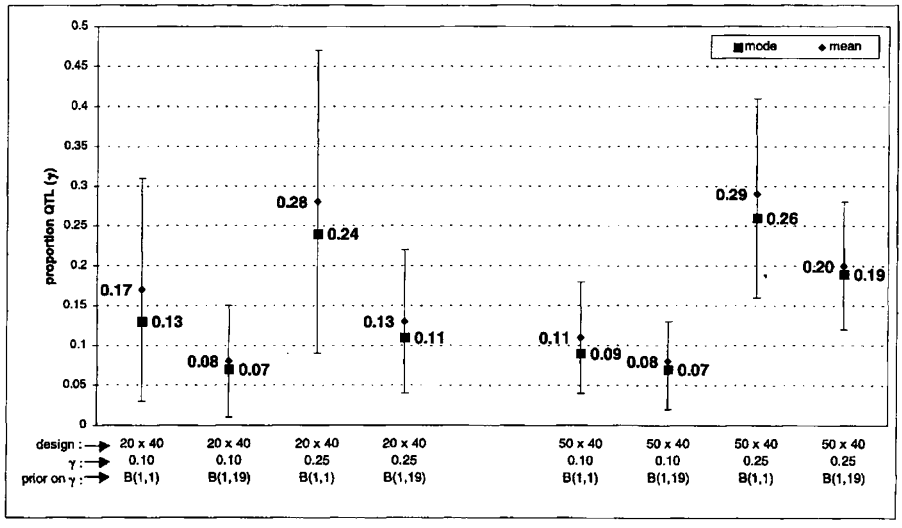

Figure 5. Estimates for posterior mode, mean and 90 % highest posterior density (HPD90) region. Estimates are averages over 25 replicates (simulation III).

5. CONCLUSIONS

We presented MCMC algorithms, using the Gibbs sampler and the MH algorithm, which facilitate Bayesian estimation of location and dispersion parameters with a RAM. The RAM proved to be superior to the FAM; RAM required much less computational time because of the greatly reduced number of location

parameters and also better mixing of the dispersion parameters. Information on individual phenotypes led to accurate estimation of both residual variance and heritability, as was similar to Van Arendonk et al. [27]. On the contrary, daughter yield deviations [28] may result in poor estimation of polygenic and residual variances [25]. The use of θ_{RT} allows a better representation of prior belief about dispersion parameters while sampling efficiency was similar to the usual θ_{VC} parameterization.

Considering ratios of variance components rather than variance components themselves in sampling procedures has been previously proposed [23]. However, our ratios can be interpreted directly and have implicit boundaries (zero and one), where Theobald et al. [23] needed a specific restriction on their ratio. The representations of prior knowledge in the two parameterizations were not equivalent and differences in posterior estimates can be expected. However, the use of vague priors (absence of prior knowledge) in the two parameterizations lead to very similar results.

In this study, position of the QTL was assumed known. Extension of the MCMC algorithm to allow estimation of QTL position has been studied and implemented [3]. Currently, the method of Bink et al. [2] to sample genotypes for a single marker is being extended to multiple markers linked to a normally distributed QTL. Then, a robust MCMC method becomes available for linkage analysis in multiple generation pedigrees allowing incomplete information on both trait phenotypes and marker genotypes.

ACKNOWLEDGEMENT

The authors wish to thank Luc Janss and George Casella for stimulating discussions and suggestions. Comments from anonymous reviewers and the editor considerably improved the paper. The first author acknowledges financial support from NWO while on research leave at Cornell University, Ithaca, NY. The financial support of Holland Genetics is gratefully acknowledged.

REFERENCES

[1] Berger J.O., Statistical Decision Theory and Bayesian Analysis, 2nd ed., Springer-Verlag, New York, NY, 1985.

[2] Bink M.C.A.M., Van Arendonk J.A.M, Quaas R.L., Breeding value estimation with incomplete marker data, Genet. Sel. Evol. 30 (1998) 45–48.

[3] Bink M.C.A.M., Janss L.L.G., Quaas R.L., Mapping a poly-allelic quantitative trait locus using simulated tempering, Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, University of New England, Armidale NSW 2351, vol. 26, Australia, 1998, pp. 277–280.

[4] Cantet R.J.C., Smith C., Reduced animal model for marker assisted selection using best linear unbiased prediction, Genet. Sel. Evol. 23 (1991) 221–233.

[5] Casella G., Berger R.L., Statistical Inferences, Duxbury press, Belmont, CA, 1990.

[6] Chib S, Greenberg E., Understanding the Metropolis Hastings algorithm, Am. Stat. 49 (1995) 327–335.

[7] Devroye L., Non-uniform Random Variate Generation, Springer-Verlag Inc., New York, 1996.

[8] Fernando R.L., Grossman M., Marker-assisted selection using best linear unbiased prediction, Genet. Sel. Evol. 21 (1989) 467–477.

[9] Gelfand A.E., Gibbs sampling (a contribution to the encyclopedia of Statistical Sciences), Technical Report, Department of Statistics, University of Connecticut, 1994.

[10] Gelfand A.E., Smith A.F.M., Sampling-based approaches to calculating marginal densities, J. Am. Statist. Assoc. 85 (1990) 398–409.

[11] Geman S., Geman D., Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, IEEE Trans. Pattern. Anal. Machine Intelligence 6 (1984) 721–741

[12] Gilks W.R., Wild P., Adaptive rejection sampling for Gibbs sampling, Appl. Stat. 41 (1992) 337–348.

[13] Gilks W.R., Best N.G., Tan K.K.C., Adaptive rejection Metropolis sampling within Gibbs sampling, Appl. Stat. 44 (1995) 455–472.

[14] Hastings W.K., Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109.

[15] Hoeschele I., Bayesian QTL mapping via the Gibbs Sampler, Proc. 5th World Congr. Genet. Appl. Livest. Prod. Guelph, University of Guelph, Canada, vol. 21, 1994, pp. 241–244.

[16] Hoeschele I., Van Raden P.M., Bayesian analysis of linkage between genetic markers and quantitative trait loci, I Prior knowledge, Theor. Appl. Genet. 85 (1993) 953–960.

[17] Janss L.L.G., Thompson R., Van Arendonk J.A.M., Application of Gibbs sampling for inference in a mixed major gene-polygenic inheritance model in animal populations, Theor. Appl. Genet. 91 (1995) 1137–1147.

[18] Metropolis, N.., Rosenbluth A.W., Rosenbluth M.N., Teller H., Teller E., Equations of state calculations by fast computing machines, J. Chem. Physics 21 (1953) 1087–1091.

[19] Quaas R.L., Pollak E.J., Mixed model methodology for farm and ranch beef cattle testing programs, J. Anim. Sci. 51 (1980) 1277–1287.

[20] Searle S.R., Matrix Algebra Useful for Statistics, John Wiley & Sons, New York, NY, 1982.

[21] Sorensen D.A., Andersen S., Gianola D., Korsgaard I., Bayesian inference in threshold models using Gibbs sampling., Genet. Sel. Evol. 27 (1995) 229–249.

[22] Tanner M.A., Wong W.H., The calculation of posterior distributions by data augmentation, J. Am. Stat. Assoc. 82 (1987) 528–540.

[23] Theobald C.M., Firat M.Z., Thompson R., Gibbs sampling, adaptive rejection sampling and robustness to prior specification for a mixed linear model, Genet. Sel. Evol. 29 (1997) 57–72.

[24] Tierney L., Markov chains for exploring posterior distributions (with discussion), Ann. Stat. 22 (1994) 1701–1762.

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

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

[27] VanArendonk J.A.M., Tier B., Bink M.C.A.M., Bovehuis H., Restricted maximum likelihood analysis between genetic markers and quantitative trait loci for a granddaughter design, J. Dairy. Sci. (1997).

[28] VanRaden P.M., Wiggans G.R., Derivation, calculation and use of national animal model information, J. Dairy Sci. 74 (1991) 2737–2746.

[29] Van Tassell C.P., Casella G., Pollak E.J., Effects of selection on estimates of variance components using Gibbs sampling and restricted maximum likelihood, J. Dairy Sci. 78 (1995) 678–692.

[30] Van Tassell C.P., VanVleck L.D., Multiple-trait Gibbs sampler for animal models: flexible programs for bayesian and likelihood-based (Co)variance component Inference, J. Anim. Sci. 74 (1996) 2586–2597.

[31] Wang C.S., Rutledge J.J., Gianola D., Marginal inferences about variance components in a mixed linear model using Gibbs sampling, Genet, Sel. Evol. 25 (1993) 41–62.

[32] Wang C.S., Quaas R.L., Pollak E.J., Bayesian analysis of calving ease scores and birth weights, Genet. Sel. Evol. 29 (1997) 117–143.

[33] Wang T., Fernando R.L., VanderBeek S., Grossman M., Van Arendonk J.A.M., Covariance between relatives for a marked quantitative trait locus, Genet. Sel. Evol. 27 (1995) 251–272.

[34] Weller J.I., Kashi Y., Soller M., Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle, J. Dairy Sci. 73 (1990) 2525–2537.

A1. APPENDIX: Full conditional densities

A1.1. Location parameters

The conditional densities for location parameters are the same with either set of dispersion parameters (θ_{VC} or θ_{RT}). When sampling genetic effects, the ratios of VC needed can be computed from either parameterization, i.e. $\alpha_{\mathrm{u}}^{-1} = \dfrac{\sigma_{\mathrm{u}}^2}{\sigma_{\mathrm{e}}^2} = \left((1-\gamma) \times \dfrac{h^2}{1-h^2}\right)$, and $\alpha_{\mathrm{v}}^{-1} = \dfrac{\sigma_{\mathrm{v}}^2}{\sigma_{\mathrm{e}}^2} = \left(\dfrac{1}{2}\gamma \times \dfrac{h^2}{1-h^2}\right)$. In this study we considered only one fixed effect, an overall mean m, for which the conditional density becomes

$$\mu|\theta_{-\mu},\ \mathbf{y} \sim N\left(\frac{1}{N}\left(\sum_{i=1}^{4}\sum_{k=1}^{n_i}\tilde{\mu}_k\right),\ \left(\sum_{i=1}^{4} n_i \sigma_{\varepsilon_i}^{-2}\right)^{-1}\right)$$

where $\tilde{\mu}_r$ equals y_k corrected for genetic effects, following the categorization in *table I*. The conditional variance of this overall mean is a weighted average over categories. Again, for phenotypes on animals in categories 1 to 3, the residual variance, $\sigma_{\varepsilon_i}^2$, contains parts of the genetic variances. The conditional density for the polygenic effect of animal j can be given as

$$u_j|\theta_{-u_j},\ \mathbf{y} \sim N(c_j/d_j,\ \sigma_{\mathrm{e}}^2/d_j)$$

where

$$c_j = \sum_{i=1}^{n_j} \tilde{y}_i + \alpha_{\mathrm{u}}\delta_j \frac{1}{2}(u_{\mathrm{S},j} + u_{\mathrm{D},j}) + \alpha_u \sum_{k \in o_{\mathrm{p}}(j)} (j)\delta_k \frac{1}{2}(u_k - \frac{1}{2}u_{\mathrm{M},k}) + \sum_{l \in o_{\mathrm{np}}(j)} \varphi_l \frac{1}{2}\left(\tilde{\tilde{y}}_l - \frac{1}{2}u_{\mathrm{M},l}\right)$$

$$d_j = n_j + \alpha_u \left(\delta_j + \sum_{k \in o_{\mathrm{p}}(j)} \frac{1}{4}\delta_k\right) + \sum_{l \in o_{\mathrm{np}}(j)} \frac{1}{4}\varphi_l$$

where $\tilde{y}_i$ is the ith phenotype for animal j, corrected for all effects, other than polygenic, $\tilde{\tilde{y}}_l$. is the average of phenotypes on non-parent l, also corrected for all effects other than polygenic, $o_{\mathrm{p}}(j)$ represents the offspring of animal j, which are parents themselves, $o_{\mathrm{np}}(j)$ represents the offspring of animal j, which are non-parents. Furthermore, $u_{\mathrm{M},k}$ is the polygenic effect of the other (if known) parent (mate of animal j) of offspring k, n_j is the number of phenotypes for animal j, δ_j, $= 1$, $4/3$, 2 when 0, 1, or 2 parents of j are identified (with no inbreeding). (δ_j^{-1} is the fraction σ_{u}^2 term ϕ_j.) Finally, φ_l is the reciprocal of the amount of variance present in the residuals of phenotypes on animal l, and can be calculated as

$$\varphi_l = \left(n_l^{-1} + \alpha_u^{-1}\delta_l^{-1} + \alpha_{\mathrm{v}}^{-1}\mathbf{1}_2^{\mathrm{T}}\mathbf{D}_l\mathbf{1}_2\right)^{-1}$$

where n_l is the number of observations on animal l, and $\mathbf{D}_l = \mathbf{I}_2 - \mathbf{Q}_l \times \mathbf{Q}_l^T$ (with no inbreeding, see also Bink et al. [2]). The conditional density for the xth QTL effect of animal j can be given as

$$v_j^x | \theta_{-v_j^x},\ \mathbf{y} \sim N(c_j^x/d_j^x,\ \sigma_{\mathrm{e}}^2/d_j^x), \quad x = 1, 2$$

where

$$c_j^x = \sum_{i=1}^{n_j} \tilde{y}_i + \alpha_v \left(dq_j^{x,1} v_{\mathrm{S},j}^1 + dq_j^{x,2} v_{\mathrm{S},j}^2 + dq_j^{x,3} v_{\mathrm{D},j}^1 + dq_j^{x4} v_{\mathrm{D},j}^2\right) + \alpha_v \sum_{k \in o_{\mathrm{p}}(j)} \left(dq_k^{1,x} v_k^1 + dq_k^{2,x} v_k^2 - dqd_k^{3,x} v_{\mathrm{M},j}^1 - dqd_k^{4,x} v_{\mathrm{M},j}^2\right) + \sum_{l \in o_{\mathrm{np}}(j)} \varphi_l \left(\tilde{\tilde{y}}_l - qdq_{\mathrm{M},l}^{x,3} v_{\mathrm{M},l}^1 - qdq_{\mathrm{M},l}^{x,4} v_{\mathrm{M},l}^2\right) - \left(n_j + \alpha_v \mathbf{D}_j^{12} + \alpha_v \sum_{k \in o_{\mathrm{p}}(j)} dqd_k^{12} + \sum_{l \in o_{\mathrm{np}}(j)} \frac{1}{2}\varphi_l \frac{1}{2}\right) v_k^{(3-x)}$$

and $d_j^x = n_j + \alpha_{\mathrm{v}}\left(\mathbf{D}_j^{xx} + \sum_{k \in o_{\mathrm{p}}(j)} dqd_k^{xx}\right) + \sum_{l \in o_{\mathrm{np}}} \frac{1}{2}\varphi_1 \frac{1}{2}$

$\tilde{y}_i$ is the ith phenotype for animal j, corrected for all effects other than QTL, $\widetilde{\widetilde{y}}_l$. is the average of phenotypes on non-parent l, also corrected for all effects other than QTL, $dq_j^{x,1}$ is the first element of the xth row of $\mathbf{D}_j^{-1}\mathbf{Q}_j^{\mathrm{T}}$ for animal j, and corrects for the covariance at the QTL between parent and offspring. Similarly, $dqd_j^{x,1}$ is the first element of the xth row of $\mathbf{Q}_j\mathbf{D}_j^{-1}\mathbf{Q}_j^{\mathrm{T}}$ for animal j, and corrects for the covariance between parent and the mate belonging to a particular offspring of that parent j.

A1.2. Dispersion parameters

In the RAM, the residuals (**e**) have different variances over the categories of animals (*table I*). Hence, conditional densities for VC in θ_{VC} are not standard densities. For example, when deriving the full conditional density for σ_{e}^2, the term $\omega_i(\sigma_{\mathrm{u}}^2 + 2\sigma_{\mathrm{v}}^2)$ is known in the likelihood part of the joint posterior density (13). It can thus be treated as a constant, but it does not drop out of the equation. With θ_{RT}, the conditional density of σ_{e}^2 is standard, but those for h^2 and γ are not. With θ_{VC}, the conditional density of variance component x, for $x = \mathbf{e}$, $\mathbf{u}$ or $\mathbf{v}$, is

$$f(\sigma_x^2|\theta_{VC,-\sigma_x^2}, \mathbf{y})$$
$$= \mathbf{p}(\sigma_x^2 \times \prod_{i=1}^{4}\left[\left(\tau(\omega_i)\sigma_{\mathrm{e}}^2\right)^{-0.5n_i} \times \exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{n_i}\mathbf{e}_k^2 \Big/ \left(\sigma_{\mathrm{e}}^2\tau(\omega)\right)\right)\right\}\right] \times \mathbf{q}(x)$$

where

$$\tau(\omega_i) = 1 + \omega_i(\sigma_{\mathrm{u}}^2 + 2\sigma_{\mathrm{v}}^2)/\sigma_{\mathrm{e}}^2 = 1 + \omega_i h^2/(1-h^2)$$
$$\mathbf{p}(\sigma_x^2) = (\sigma_x^2)^{-\alpha_x - 1}\exp\left\{\frac{-1}{\beta_x\sigma_x^2}\right\}$$

and

$$q(x) = \begin{cases} 1 & \text{if } x = e \\ (\sigma_{\mathrm{u}}^2)^{-0.5q} \times \exp\left\{-\frac{1}{2}(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}) \times \frac{1}{\sigma_{\mathrm{u}}^2}\right\} & \text{if } x = u \\ (\sigma_{\mathrm{v}}^2)^{-0.5(2q)} \times \exp\left\{-\frac{1}{2}(\mathbf{v}^T\mathbf{G}^{-1}\mathbf{v}) \times \frac{1}{\sigma_{\mathrm{v}}^2}\right\} & \text{if } x = v \end{cases}$$

With θ_{RT}, the conditional density for σ_{e}^2 is an $IG(r,s)$ distribution with

$$r = \left[\alpha_{\mathrm{e}} + \frac{1}{2}N + \frac{1}{2}q + \frac{1}{2}(2q)\right]$$

$$s = \left[\frac{1}{\beta_{\mathrm{e}}} + \frac{1}{2}\left(\sum_{i=1}^{4}\sum_{k=1}^{n_i}\frac{e_k^2}{\tau(\omega_i)}\right) + \frac{1}{2}\frac{\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}}{(1-\gamma) \times \frac{h^2}{1-h^2}} + \frac{1}{2}\frac{\mathbf{v}^T\mathbf{G}^{-1}\mathbf{v}}{0.5\gamma \times \frac{h^2}{1-h^2}}\right]$$

where N is the total number of phenotypes

$$f\left(h^2|\theta_{RT,-h^2},\ \mathbf{y}\right) \propto (h^2)^{\alpha_{h^2}-1}(1-h^2)^{\beta_{h^2}-1}$$
$$\times \prod_{i=1}^{4}\left[\tau(\omega_i)^{-0.5n_i} \times \exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{n_i} e_k^2/\left(\tau(\omega_i)\sigma_e^2\right)\right)\right\}\right]$$
$$\times \left(\frac{h^2}{1-h^2}\right)^{-0.5(q+2q)} \times \exp\left\{-\frac{1}{2}[(\mathbf{u}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{u}/(1-\gamma)) + (\mathbf{v}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{v}/0.5\gamma)] \times \frac{1-h^2}{h^2\sigma_e^2}\right\}$$

where $\tau(\omega_i) = 1 + \omega_i h^2/(1-h^2)$.

$$f(\gamma|\theta_{RT,-\gamma},\mathbf{y}) \propto (\gamma)^{\alpha_\gamma-1}(1-\gamma)^{\beta_\gamma-1}(1-\gamma)^{-0.5(q)} \times (\gamma)^{-0.5(2q)}$$
$$\times \exp\left\{-\frac{1}{2}\left[(\mathbf{u}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{u}/(1-\gamma)) + (\mathbf{v}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{v}/0.5\gamma)\right] \times \frac{1-h^2}{h^2\sigma_e^2}\right\}$$