Scaling to account for heterogeneous variances in a Bayesian analysis of broiler quantitative trait loci

J. B. C. H. M. van Kaam*, M. C. A. M. Bink†, H. Bovenhuis*2, and R. L. Quaas‡

*Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, 6700 AH, Wageningen, The Netherlands; †Centre for Biometry, Plant Research International, Wageningen, The Netherlands and Department of Genetics & Reproduction, Institute for Animal Science and Health, Lelystad, The Netherlands; and ‡Department of Animal Science, Faculty of Agriculture and Life Sciences, Cornell University, Ithaca, NY

ABSTRACT: A Bayesian method for QTL analysis that is capable of accounting for heterogeneity of variance between sexes, is introduced. The Bayesian method uses a parsimonious model that includes scaling parameters for polygenic and QTL allelic effects per sex. Furthermore, the method employs a reduced animal model to increase computational efficiency. Markov Chain Monte Carlo techniques were applied to obtain estimates of genetic parameters. In comparison with previous regression analyses, the Bayesian method 1) estimates dispersion parameters and polygenic effects, 2) uses individual observations instead of offspring averages, and 3) estimates fixed effect levels and covariates and heterogeneity of variance between sexes simultaneously with other parameters, taking uncertainties fully into account. Broiler data collected in a feed efficiency and a carcass experiment were used to illustrate QTL analysis based on the Bayesian method. The experiments were conducted in a population consisting of 10 full-sib families of a cross between two broiler lines. Microsatellite genotypes were determined on generation 1 and 2 animals and phenotypes were collected on third-generation offspring from mating members from different families. Chromosomal regions that seemed to contain a QTL in previous regression analyses and showed heterogeneity of variance were chosen. Traits analyzed in the feed efficiency experiment were BW at 48 d and growth, feed intake, and feed intake corrected for BW between 23 and 48 d. In the carcass experiment, carcass percentage was analyzed. The Bayesian method was successful in finding QTL in all regions previously detected.

Key Words: Bayesian Theory, Chickens, Genetic Markers, Heterogeneity, Monte Carlo Method, Quantitative Trait Loci

©2002 American Society of Animal Science. All rights reserved.


Introduction

In recent years, the availability of genetic markers has increased rapidly (Rohrer et al., 1996; Groenen et al., 2000). Marker information enables mapping QTL by reconstruction of the transmission of chromosomal segments from parents to offspring. Statistical QTL mapping methods (for review see Bovenhuis et al., 1997; Hoeschele et al., 1997) differ in their computational requirements and ability to handle different population structures. Simple methods such as regression interval mapping are very suitable for initial genomewide analyses. Bayesian analysis, facilitated by sampling from conditional parameter distributions via Marker Chain Monte Carlo (MCMC) techniques, is computationally demanding but can take full account of the uncertainty associated with all the unknown parameters (Wang, 1998). When applied to an animal model including polygenic and QTL effects with relationship matrices, a Bayesian analysis is not limited to a specific pedigree structure and can accommodate partly missing marker genotypes (Bink and Van Arendonk, 1999).

In previous studies, Van Kaam et al. (1998, 1999a,b) performed whole genome scans and identified QTL affecting growth, feed efficiency, and carcass traits in broilers using regression interval mapping. This approach accounted only for the most likely haplotype configuration, did not use polygenic relations, and was limited for usage in complex populations because only genotypes from two generations are used. Furthermore, preadjust-
of offspring observations for fixed effects, heterogeneity of variance between sexes, and parental mate contributions is required. In a preadjustment for heterogeneity of variance, it is not possible to distinguish polygenic, QTL, and environmental variance. In the present study, a method was developed that simultaneously handles fixed, polygenic, and QTL effects while accounting for heterogeneity of variance between sexes and uncertainties. This method is applied to fixed positions in chromosomal regions where QTL were previously found.

Materials and Methods

Animal Model

Fernando and Grossman (1989) extended the animal model by including normally distributed QTL effects in addition to a polygenic effect. Some other studies applying an animal model with polygenic and QTL effects were done by Uimari et al. (1996) and Meuwissen and Goddard (1997). The Fernando and Grossman model can be represented as follows:

\[
y|\mathbf{b}, \mathbf{u}, \mathbf{v} \sim \mathcal{N}(\mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{v}, \mathbf{L}\sigma^2)
\]

with

\[
\begin{bmatrix}
\mathbf{u} \\
\mathbf{v}
\end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{A} \sigma^2_u & 0 \\
0 & \mathbf{G}_k\sigma^2_v \end{bmatrix}\right)
\]

where \(\mathbf{y}\) is a \(n\)-vector of phenotypes, \(\mathbf{X}\) is a \(n \times p\) incidence matrix relating fixed effect levels and covariates to phenotypes, \(\mathbf{b}\) is a \(p\)-vector of fixed effect levels and covariates, \(\mathbf{Z}\) is a \(n \times q\) incidence matrix relating individuals to phenotypes, \(\mathbf{u}\) is a \(q\)-vector of random additive polygenic effects, \(\mathbf{W}\) is a \(n \times 2q\) incidence matrix relating QTL allelic effects to phenotypes, \(\mathbf{v}\) is a \(2q\)-vector of random additive QTL allelic effects, \(\mathbf{A}\) is the additive genetic relationship matrix, \(\sigma^2_u\) is the polygenic variance excluding the QTL, \(\mathbf{G}_k\) is the gametic relationship matrix for the QTL and depends on the QTL position \(k\) and the marker information, and \(\sigma^2_v\) is the additive variance of the QTL allelic effects. Here, polygenic and QTL variances are assumed to be independent. The same error variance, \(\sigma^2_e\), is applied for all observations, hence error terms are assumed to be uncorrelated with homogeneous variance.

In the present experiments, heterogeneity of variance between sexes occurs (e.g., BW-related traits) (Van Kaam et al., 1998). For example, the 95\% confidence intervals for the phenotypic variance of BW at 48 d were [55,018; 65,502] for males and [39,116; 46,471] for females. Van Kaam et al. (1999a,b) previously reported additive genetic correlations between sexes ranging from 0.87 to 1. Therefore, we assume that the same genes are responsible for these traits in both sexes and we postulate that the genetic part of the heterogeneity is due to differences in the magnitude of allelic effects between sexes. Hence, heterogeneity can be either due to different polygenic effects, QTL effects, or otherwise fixed or random environmental effects. For the genetic effects, heterogeneity is modeled with the introduction of scale parameters (Quaas et al., 1989). Separate scale parameters per sex are used for polygenic and QTL allelic effects. Furthermore, separate fixed effects and error variances are modeled per sex. This leads to the scaled model:

\[
y_s|\mathbf{b}_s, \mathbf{u}, \mathbf{v} \sim \mathcal{N}(\mathbf{X}\mathbf{b}_s + c_s\mathbf{Z}\mathbf{u} + d_s\mathbf{W}\mathbf{v}, \mathbf{L}\sigma^2_s) \text{ for } s = m, f
\]

with

\[
\begin{bmatrix}
\mathbf{u} \\
\mathbf{v}
\end{bmatrix} \sim \mathcal{N}(0, \begin{bmatrix} \mathbf{A} & \mathbf{0} \\
\mathbf{0} & \mathbf{G}_k \end{bmatrix})
\]

where \(c_s\) and \(d_s\) represent scale parameters for the polygenic and QTL allelic effects, respectively, and subscript \(s\) indicates sex: male (\(m\)) or female (\(f\)). In the scaled model, the variances of the random genetic effects are fixed, because otherwise the scale parameters and these variance components would both be measuring the same dispersion and not both be identifiable. The solution, taken here, is to fix \(\sigma^2_u\) and \(\sigma^2_v\) to one and hence \(\mathbf{u}\) and \(\mathbf{v}\) have a standard normal distribution. Rather than a single polygenic variance \(\sigma^2_u\), as in the homoskedastic case, we now have \(c_m^2\) and \(c_f^2\), depending on the sex in which genes are expressed. Likewise, we have \(d_m^2\) and \(d_f^2\) for variances of QTL allelic effects. The total additive genetic variance equals the polygenic variance and twice the QTL allelic variance. A scale parameter can be interpreted as a standard deviation, but in the model equation it is a regression coefficient. Regression coefficients typically have normal conjugate priors like other mean effects. Here a left-truncated normal prior is used to assure non-negativity for the scale parameters:

\[
c_s \sim \text{TN}(\mu_{c_s}, \sigma^2_{c_s}) \text{ with } c_s \geq 0
\]

\[
d_s \sim \text{TN}(\mu_{d_s}, \sigma^2_{d_s}) \text{ with } d_s \geq 0
\]

The properness of the posterior follows from the truncated normal prior and the normal sample density both being proper (Hobert and Casella, 1996). The advantage of the model specification using scale parameters instead of variances is that it shows more clearly that heterogeneity is considered as a scale effect; furthermore, parameters, which are not in the quadratic form, are easier to interpret. Uniform priors for fixed effects are assumed. In the present case, uncorrelated genetic effects and uncorrelated error terms with homogeneous variance within sex are assumed. For the genetic effects, normally distributed priors are taken; hence, for the polygenic effects \(\mathbf{u} \sim \mathcal{N}(0, \mathbf{A})\) and for the QTL allelic effects \(\mathbf{v} \sim \mathcal{N}(0, \mathbf{G}_k)\). Inverted Gamma distributions with predefined hyperparameters \(\alpha\) and \(\lambda\) are used to represent prior knowledge on the error variances as \(\sigma^2_e - IG(\alpha, \lambda)\).
MCMC Algorithm

The solutions of the model are obtained using MCMC techniques, which enable sampling from the posterior distribution of parameters. A reduced animal model (RAM) was used to obtain solutions more efficiently because polygenic effects for nonparents and QTL allelic effects for ungenotyped nonparents do not have to be sampled (Cantet and Smith, 1991; Bink et al., 1998). In Appendix 1, the full conditional distributions of the fixed and random genetic effects are presented. Fixed effect levels and covariates, random polygenic and QTL allelic effects, and haplotypes are sampled using Gibbs sampling (Bink et al., 1998). The RAM residuals of nonparents consist of an error term and the Mendelian parts of the additive genetic variance depending on the RAM category. With the scaled RAM, the full conditionals for these parameters are not standard distributions because the scale parameters appear in both the means and variance of \( p(y_{ij}|b,u,v,c,d,s,\sigma^2) \). Therefore, the conditional distributions of the dispersion parameters do not have a simple form to facilitate Gibbs sampling. Hence, Metropolis-Hastings is used to sample scale parameters and error variances. The Metropolis-Hastings algorithm is presented in Appendix 2.

The positions of the markers and QTL are not sampled but fixed. The QTL position \( k \) is modeled by including information from flanking markers \( m \) in the computation of the inverse of the gametic relationship matrix \( G_{ij} \) (Wang et al., 1995). Marker information is described in terms of the allelic constitution of the chromosomal homologues of the founders and identity by descent values for all nonfounders (Jansen et al., 1998; Bink and Van Arendonk, 1999).

The most important limitations of the method are that the QTL position is fixed and that founder marker genotypes are needed.

Experimental Population

A three-generation population was created for the purpose of QTL detection, following recommendations of Van der Beek et al. (1995). Founder animals, parents, offspring, and grand-offspring are indicated as generation 0, 1, 2, and 3 animals or \( G_0, G_1, G_2 \), and \( G_3 \) animals, respectively. In the three-generation design, \( G_1 \) and \( G_2 \) animals were typed for genetic markers and phenotypic observations were collected on different hatches of \( G_3 \) animals. Genotypes for some markers, however, were created, which on average produced 45.1 \( G_2 \) full sibs. \( G_2 \) animals were mated with several \( G_2 \) animals from different families to produce nine \( G_3 \) animals on average. In the analyses \( G_1, G_3, \) and \( G_3 \) animals were included and \( G_0 \) animals were omitted, because they were not genotyped and our Bayesian method requires known marker genotypes for base animals.

Traits

In this paper we limit the analysis to those traits that showed heterogeneity of variance between sexes and suggestive significance for the presence of QTL in previous analyses (Van Kaam et al., 1999a,b). Phenotypes (n = 2,049) analyzed in this study included body weight at 48 d (BW48; \( h^2 \) in males = 0.28 and \( h^2 \) in females = 0.33) and growth (GAIN; \( h^2 \) in males = 0.23 and \( h^2 \) in females = 0.19), feed intake (FIFA; \( h^2 \) in males = 0.25 and \( h^2 \) in females = 0.39), and feed intake adjusted for BW (FIFW; \( h^2 \) in males = 0.36 and \( h^2 \) in females = 0.39); all three were measured between 23 and 48 d in a feed efficiency experiment (Van Kaam et al., 1999b). Carcass percentage (CP; \( h^2 \) in males = 0.43 and \( h^2 \) in females = 0.52) (n = 1,953) measured in a carcass experiment was also analyzed, although Van Kaam et al. (1999a) found a low heterogeneity between sexes, because it is derived from body and carcass weight, which do show heterogeneity. A few outlying observations more than 3 SD from the mean of the hatch were removed from the analysis.

Fixed effect subclasses for BW48, GAIN, FIFA, and FIFW were based on the location of the animal’s cage within the building and an interaction between the hatch of the dam and the hatch of the animal. For FIFW, the observations of FIFA were used with BW23 and BW48 as covariates. For CP, an interaction between hatch of the dam, hatch of the offspring, and the day of measuring carcass weight was included as a fixed effect. Because carcass weight was measured on 2 d in one hatch and slaughtering was on one day, dehydration might have had an influence on the measurement within hatch, and therefore, the day of measuring was included in the interaction term. All fixed effects were estimated within sex.

Marker Data

Genotypes for microsatellite markers were determined using DNA derived from blood samples from all 20 \( G_1 \) and 451 \( G_2 \) animals. Marker alleles were recorded in base-pair units. Marker data used in these analyses were a subset of the marker data used for creating the linkage map (Groenen et al., 1998). Only five chromosomal regions that showed heterogeneity of variance between sexes and suggestive significance for the presence of a QTL in previous analyses (Van Kaam et al., 1999a,b) were selected for further analysis. Marker alleles were determined in all 10 families for most of the markers in these regions. Genotypes for some markers, however, were only collected in four families. More details on the
Table 1. Population structure with numbers of animals used in the analyses and types of observations collected

<table>
<thead>
<tr>
<th>Generation</th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>G₀</td>
<td>14</td>
<td>14</td>
<td>28</td>
<td>—</td>
</tr>
<tr>
<td>G₁</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>Genotypes</td>
</tr>
<tr>
<td>G₂</td>
<td>172</td>
<td>279</td>
<td>451</td>
<td>Genotypes</td>
</tr>
<tr>
<td>G₃</td>
<td>1,012</td>
<td>1,037</td>
<td>2,049</td>
<td>Phenotypes</td>
</tr>
<tr>
<td>G₄</td>
<td>969</td>
<td>984</td>
<td>1,953</td>
<td>Phenotypes</td>
</tr>
</tbody>
</table>

*Numbers exclude outliers and missing values.

**G₀, etc.** = Generation 0, etc.

*Male and female G₀ animals are from different lines; G₀ animals were not included in the analyses because marker genotypes were unknown.

BW₄₈ = body weight at 48 d; FIFA = feed intake in a fixed age interval; FIFW = feed intake in a fixed weight interval; GAIN = growth between 23 and 48 d; CP = carcass percentage.

regions analyzed are given in Figure 1. A minimum marker spacing of about 2 cM was aimed at, except for the most lateral markers, which were used to increase informativity at the ends of the map. On all analyzed regions, all 20 parents were informative, except on linkage group WAU26, where four parents were uninformative. The genotypes of two markers with the same position, MCW0023 and ADL0183, were combined into a single haplotyped locus. The same was done for LEI0084 and MCW0181.

**MCMC and Prior Distribution Settings**

For all chromosomal regions of interest, several independent QTL analyses were each based on a single chain of 2,000,000 cycles after 1,000 cycles burn-in time. A
the variance on the polygenic scale parameter, \( \mu \), and QTL parameters at zero, hence additive genetic variance with the mode of the QTL scale variance explained by the putative QTL is 20\% of the resulting in observed variance without adjustment for fixed effects, ing that 1) the expected residual variance is 40\% of the stored. Dispersion parameters after the burn-in time were very time-demanding to sample. All samples of the haplotypes were sampled every 50th cycle, because they are very time-demanding to sample. All samples of the dispersion parameters after the burn-in time were stored.

Priors for dispersion parameters were chosen assuming that 1) the expected residual variance is 40\% of the observed variance without adjustment for fixed effects, resulting in \( \lambda \); 2) the heritability is 0.3; 3) the expected variance explained by the putative QTL is 20\% of the additive genetic variance with the mode of the QTL scale parameters at zero, hence \( \mu_d = 0 \) without truncation; 4) the variance on the polygenic scale parameter, \( \sigma^2_p \), is 0.09 \( \times \) the expected polygenic variance; and 5) there is no heterogeneity of variance between sexes (i.e., the same priors were used for males and females). Absence of heterogeneity was assumed so that the heterogeneity in the posterior only comes from the data. Using the first two assumptions, the additive genetic variance can be calculated. With the third assumption, the additive genetic variance can be divided over the polygenic and QTL variance. Then \( \sigma^2_d \) follows from the expected QTL variance and \( \mu_d \) is obtained from a small simulation. The \( \alpha \) hyperparameter of the inverted Gamma prior for the error variances was 2.000001 in all cases, to obtain an inverse chi-square. The settings for the prior distributions for the scale parameters as well as the \( \lambda \) hyperparameter of the inverted Gamma prior for the error variances are shown in Table 2. In analyses of models with and without QTL the total additive genetic variance was assumed equal. Therefore, the polygenic variance is larger in models without QTL. The putative QTL variance explaining 20\% of the additive genetic variance was chosen because the QTL effects found in the previous studies (Van Kaam et al., 1999a,b) were quite large and the number of phenotypes made it hard to find smaller QTL. The starting values chosen for the dispersion parameters were the same values as given for \( \mu_e, \mu_d, \), and \( \lambda \) in Table 2. The short burn-in time was chosen because the dispersion parameters were started at their prior expectations. The acceptance probabilities for polygenic and QTL scale parameters and error variances were 0.83, 0.74, and 0.61.

**Results**

**Heterogeneity of Variance Between Sexes**

Table 3 shows the posterior means for the estimated heritabilities including the QTL, the QTL proportion of the total genetic variance, and the phenotypic variance. Results are shown for the marker bracket most likely containing the QTL (i.e., the bracket with the largest QTL effect). Heritabilities were larger for models that included a QTL. Differences in phenotypic variances between males and females were found for all traits. For most of the traits, the phenotypic variance in males was larger than in females, except for carcass percentage, in which female phenotypic variance was larger. The heterogeneity of variance was most pronounced for GAIN: male phenotypic variance was 1.5 times the female phenotypic variance. The coefficient of variation for males and females was 0.131 and 0.146, respectively. The estimated male and female heritabilities for most traits were in the same order, suggesting that this

---

**Table 2. Values of the left-truncated normal priors for scale parameters and inverted Gamma priors for the error variances**

<table>
<thead>
<tr>
<th>Trait</th>
<th>( \mu_e )</th>
<th>( \sigma^2_e )</th>
<th>( \mu_d )</th>
<th>( \sigma^2_d )</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model including QTL</td>
<td></td>
<td>Model without QTL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAIN</td>
<td>94</td>
<td>876</td>
<td>0</td>
<td>1,217</td>
<td>28,400</td>
</tr>
<tr>
<td>BW48</td>
<td>113</td>
<td>1,271</td>
<td>0</td>
<td>1,766</td>
<td>41,200</td>
</tr>
<tr>
<td>FIFA</td>
<td>165</td>
<td>2,703</td>
<td>0</td>
<td>3,754</td>
<td>87,600</td>
</tr>
<tr>
<td>FIFW</td>
<td>79</td>
<td>617</td>
<td>0</td>
<td>857</td>
<td>20,000</td>
</tr>
<tr>
<td>CP</td>
<td>0.64</td>
<td>0.04</td>
<td>0.00</td>
<td>0.06</td>
<td>1.3</td>
</tr>
</tbody>
</table>

*Indicated per trait are the prior values for left-truncated normal priors of the polygenic scale parameters and the QTL allelic scale parameters and inverted Gamma priors of the error variances. The values of \( \mu_e \) and \( \sigma^2_e \) presented are before truncation. The truncation point for the scale parameters was at zero. The \( \alpha \) hyperparameter of the inverted Gamma prior was 2.000001 in all cases. For both sexes, the same prior values were used.

**bw48** = body weight at 48 d; **FIFA** = feed intake in a fixed age interval; **FIFW** = feed intake in a fixed weight interval; **GAIN** = growth between 23 and 48 d; **CP** = carcass percentage.
Table 3. Posterior means of the heritability, proportion QTL variance of the total genetic variance, and phenotypic variances in the most likely marker bracket using a model with a QTL and a model without QTL.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Interval</th>
<th>$h^2_{m}$</th>
<th>$h^2_{f}$</th>
<th>$\gamma_{m}$</th>
<th>$\gamma_{f}$</th>
<th>$\sigma^2_{m}$</th>
<th>$\sigma^2_{f}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAIN</td>
<td>LEI0071-MCW0101</td>
<td>0.25</td>
<td>0.23</td>
<td>0.46</td>
<td>0.21</td>
<td>42,738</td>
<td>27,960</td>
</tr>
<tr>
<td>GAIN</td>
<td>No QTL</td>
<td>0.21</td>
<td>0.21</td>
<td>0.00</td>
<td>0.00</td>
<td>42,388</td>
<td>27,729</td>
</tr>
<tr>
<td>BW48</td>
<td>LEI0071-MCW0101</td>
<td>0.30</td>
<td>0.28</td>
<td>0.26</td>
<td>0.14</td>
<td>59,734</td>
<td>42,318</td>
</tr>
<tr>
<td>BW48</td>
<td>No QTL</td>
<td>0.27</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
<td>59,266</td>
<td>42,087</td>
</tr>
<tr>
<td>FIFA</td>
<td>MCW0058-LEI0071</td>
<td>0.30</td>
<td>0.33</td>
<td>0.40</td>
<td>0.13</td>
<td>144,050</td>
<td>114,879</td>
</tr>
<tr>
<td>FIFA</td>
<td>ADL0194-MCW0085</td>
<td>0.27</td>
<td>0.34</td>
<td>0.21</td>
<td>0.21</td>
<td>144,050</td>
<td>114,879</td>
</tr>
<tr>
<td>FIFA</td>
<td>ADL0262-MCW0165</td>
<td>0.27</td>
<td>0.36</td>
<td>0.23</td>
<td>0.21</td>
<td>144,050</td>
<td>114,879</td>
</tr>
<tr>
<td>FIFA</td>
<td>No QTL</td>
<td>0.25</td>
<td>0.32</td>
<td>0.00</td>
<td>0.00</td>
<td>142,789</td>
<td>114,078</td>
</tr>
<tr>
<td>FIFW</td>
<td>ADL0343-MCW0082</td>
<td>0.40</td>
<td>0.39</td>
<td>0.33</td>
<td>0.26</td>
<td>41,377</td>
<td>28,517</td>
</tr>
<tr>
<td>FIFW</td>
<td>No QTL</td>
<td>0.36</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
<td>41,377</td>
<td>28,517</td>
</tr>
<tr>
<td>CP</td>
<td>LEI0079-MCW0177</td>
<td>0.34</td>
<td>0.37</td>
<td>0.16</td>
<td>0.26</td>
<td>41,377</td>
<td>28,517</td>
</tr>
<tr>
<td>CP</td>
<td>No QTL</td>
<td>0.31</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
<td>41,377</td>
<td>28,517</td>
</tr>
</tbody>
</table>

$^a$Indicated per trait are two analyses, one showing the most likely marker bracket with a model containing a QTL and one with a model without a QTL. For each analysis, the heritability, proportion QTL variance of the total genetic variance and phenotypic variances in males and females are shown. For heritabilities and proportions QTL the HPD95 is shown between brackets.

$^b$BW48 = body weight at 48 d; FIFA = feed intake in a fixed age interval; FIFW = feed intake in a fixed weight interval; GAIN = growth between 23 and 48 d; CP = carcass percentage.

$^c$Male and female proportions QTL variance of the total genetic variance are calculated as $h^2 = (c^2_m + 2d^2_f)/(c^2_m + 2d^2_f + 2c^2_f)$ and $h^2 = (c^2_f + 2d^2_f)/(c^2_f + 2d^2_f + 2c^2_f)$.

heterogeneity is to the same extent due to differences in environmental as well as additive genetic variances. For FIFA, additive genetic variances were similar in males and females.

The polygenic variance, which can be derived from Table 3, shows heterogeneity most clearly for GAIN, FIFW, and CP. The QTL variance, which also follows from Table 3, shows heterogeneity for the QTL in the region MCW0058-MCW0101 affecting GAIN, BW48, and FIFA and for the QTL in the interval ADL0343-MCW0082 affecting FIFW. The QTL affecting FIFW in the intervals ADL0194-MCW0085 and ADL0262-MCW0165 and CP in the interval LEI0079-MCW0177 seem to have a similar effect on both sexes.

Presence of QTL

The QTL analyses show evidence for the presence of a QTL in each of the nine regions where a QTL was found in the previous regression analyses. A QTL is assumed present if a value of zero is not in the 95% Highest Posterior Density (HPD95) region for the QTL scale parameter (i.e., the QTL variance differs significantly from zero). In 8 out of 32 marker brackets, a QTL was found that affected observations in only one sex, and in 18 marker brackets both sexes were affected. The most likely marker brackets reported by Van Kaam et al. (1998, 1999a,b) and analyzed here as well, were 1) UMA1.107-MCW0058 for FIFA and GAIN on Chromosome 1; 2) MCW0058-LEI0071 for BW48 on Chromosome 1; 3) MCW0082-MCW0341 for FIFW on Chromosome 2; 4) MCW0085-LEI0122 for FIFA on Chromosome 4; 5) ADL0289-ADL0262 for FIFA on linkage group WAU26; and 6) ADL0183-LEI0079 for CP on Chromosome 1. With the Bayesian method these marker brackets always contained a significant QTL except for GAIN expressed in males and BW48 expressed in females. In several cases, a QTL seemed present in one or two of the flanking marker brackets. Therefore, we cannot be certain that the most likely marker bracket is the actual bracket containing the QTL. As shown before (Van Kaam et al., 1999b), BW48, GAIN and FIFA are correlated, and therefore the same QTL might affect these three traits. For these traits, the same chromosomal region was analyzed, which influences the consistency of the results.

Scale Parameters

In Figure 2, an example of the marginal posterior densities of male QTL scale parameters is given. These densities are the result of the analyses of BW48. In Figure 3, the marginal posterior densities of female
Figure 2. Marginal posterior densities of the male QTL scale parameters obtained in five analyses of body weight at 48 d in consecutive marker brackets on Chromosome 1. The graph was made using 101 bins.

QTL scale parameters obtained in the same analyses are given. Each of these densities is based on approximately 150 to 400 effective samples (Sorensen et al., 1995). The lag-one serial correlation of the QTL scale parameters was between 0.97 and 0.98. All other dispersion parameters usually had a larger number of effective samples. The densities shift due to information from different markers. The pattern of the densities shows that the closer to marker bracket LEI0071-MCW0101 the further the densities shift away from zero. This provides an indication that the most likely marker bracket for the position of a QTL is the bracket LEI0071-MCW0101. The pattern for this trait is similar for the male and female QTL scale parameters. The densities of the male QTL scale parameters, however, are further away from zero and hence the QTL effect tends to be larger in males than in females.

Influence of Priors

In order to obtain an idea of the influence of the settings of the prior for the QTL scale parameters, two

Figure 3. Marginal posterior densities of the female QTL scale parameters obtained in five analyses of body weight at 48 d in consecutive marker brackets on Chromosome 1. The graph was made using 101 bins.
additional analyses were done using different settings. In these settings, a QTL explaining 5% or 10%, respectively, instead of 20% of the additive genetic variance was assumed. The total additive genetic variance was the same, hence the polygenic variance was larger. These settings were $TN(123,1506)$ and $TN(120,1434)$, respectively, instead of $TN(113,1271)$ for the polygenic scale parameters. For the QTL scale parameters the settings were $TN(0,441)$ and $TN(0,883)$, respectively, instead of $TN(0,1766)$. A comparison of the densities of the male QTL scale parameter is shown in Figure 4, which gives the prior distributions reflecting a proportion QTL variance of 5%, 10%, and 20% and the posterior distributions obtained using these two priors. With a prior of 10%, the posterior mean of the QTL variance diminished with 31% compared to a prior of 20%. The posterior mean of the proportion of the additive genetic variance explained by the QTL diminished from 26% to 19%. With a prior of 5%, the posterior mean of the QTL variance diminished with 53% compared to a prior of 20%. The posterior mean of the proportion of the additive genetic variance explained by the QTL diminished further to 15%.

Discussion

Method of Analysis

Advantages of the Bayesian method as compared to the regression analysis of Van Kaam et al. (1998, 1999a,b) are that 1) all parameters except recombination rates and QTL position are sampled simultaneously, taking uncertainty into account; 2) an animal model that included fixed and polygenic effects and polygenic and gametic relationships matrices is used; 3) heterogeneity of variances between sexes is accounted for simultaneously by scaling; and 4) dispersion parameters are estimated for all random terms in the model. Because fixed-effect levels and covariates and heterogeneity of variance can be handled by the Bayesian method, individual observations instead of offspring averages can be used. A polygenic component is part of the model instead of the family effect, as in the regression analysis. For the analysis with the regression interval mapping procedure, approximations were needed in the adjustment for contributions of the parental mates to phenotypes. Use of an animal model in a Bayesian analysis offers the opportunity to exploit all relationships through relationship matrices, which abandons the need for this adjustment. The advantage of accounting for heterogeneity by scaling is that it hardly increases the computational needs (Quaas et al., 1989) because the number of parameters increases only by using fixed effects per sex and by adding two dispersion parameters for genetic effects and one for the error variance.

The current Bayesian method requires marker genotypes for all base parents. Hence, the method does not sample genotypes for ungenotyped base animals. In the experimental population, no genotypes were collected on $G_0$ animals and, therefore, this generation was excluded from the analyses. The ability to include these animals would improve the power. An important limitation of the current approach is that QTL position is fixed at certain chromosomal locations instead of being treated as an unknown parameter that is being estimated. Therefore, several analyses are required to obtain a rough estimate of the most likely QTL position.
If the actual QTL position differs from the fixed QTL position then the estimated variance explained by the QTL will be underestimated. The method could easily be extended to a multiple QTL method.

The QTL effects are assumed to be independent from the polygenic effects in our approach. This assumption does not hold and will result in biased estimates for QTL positioned on the same chromosome as the polygenic effects or in the case of linkage disequilibrium between the linkage group containing the QTL and other linkage groups (Farnir et al., 2000). Interactions between genes located on the sex chromosomes and autosomal chromosomes are potential causes of heterogeneity of variance between sexes, which are not considered in this study. The current method uses only additive polygenic, additive QTL, and residual terms as sources of heterogeneity.

**Scaled Model**

Biologically there is just one genetic constitution per animal and one genetic variation in a population; only the expression of the genes in both sexes differs. The scaled model is similar in that it assumes one genetic variation, one polygenic effect per animal, and one effect per allele and scaling of the gene effects with respect to the sex of the animal in which the gene is expressed. The scaled model is similar to a normal bivariate model in which the genetic correlation between sexes is restricted to one. The scaled model, however, has the advantage that only one polygenic and two QTL allelic effects per animal are required. A bivariate approach would require two polygenic and four QTL allelic effects per animal. Hence, the scaled model is more parsimonious than a bivariate model, which improves estimability and reduces computational requirements. A disadvantage, however, is that the scaled model can only handle correlations of one, whereas a bivariate approach would allow any correlation. Assumming that the same genes influence a trait in both sexes, it seems justified having genetic correlations of unity if the direction of the effect is the same in both sexes. Especially for a QTL, which is assumed to be a single gene, the correlation between the effects in both sexes should be one.

Ignoring heterogeneity of variance between sexes by assuming homogeneous variances would result in more emphasis on the variance in males and less in females. In the present analysis, the possibility of opposite genetic effects between sexes was omitted. The scaled model, however, can accommodate this possibility by using normally distributed priors for the scale parameters instead of left-truncated normals. Then, scale parameters can have negative as well as positive values; however, in the variance structure they will always enter as nonnegative quadratic terms. A different sign between sexes would indicate opposite effects. In case both sexes have the same sign, it does not matter whether the sign is positive or negative, because the sign of the genetic effect and the scale parameter are exchangeable.

The scale parameters are expressed relative to the fixed variance of \( u \) and \( v \). This was done because it is not practical to express the effect in one sex as a ratio of the effect in the other sex. Such a ratio would lead to problems in case the effect in the sex, which is in the denominator of the ratio, would be zero.

**Estimates**

In our previous studies (Van Kaam et al., 1998, 1999a,b) the applied regression method for QTL detection used average adjusted progeny trait values as observations, which were based on genetic variances estimated before with MTDFREML (Boldman et al., 1995). Estimates from MTDFREML are modal estimates, whereas our Bayesian estimates are means. In comparison with our previously published MTDFREML estimates, total phenotypic variances agree closely, with maximum differences of 6%. This is not surprising, because priors were based on values obtained in the previous analysis. Heritabilities, however, are different from those obtained previously. In the current study, heritabilities in males and females are more similar than in previous results. The main difference between the analyses is the inclusion of a maternal genetic effect in the previous model, except for CP; furthermore, the previous study did not contain a QTL in the model. The posterior means of the heritabilities for CP (0.51 and 0.36) were substantially lower compared with our previous REML results (0.43 and 0.52). This can be caused by the prior assumption for heritability of 0.30. It is also possible that the previous maximum likelihood estimates were not in the global maximum. Estimated posterior means of the QTL variances are between 13 and 46% of the posterior means of the total genetic variance, which seems quite large. The QTL scale parameters depend strongly on the position that is assumed, as can be seen in Figures 2 and 3. Large QTL variances can possibly be caused by using a normal distribution for QTL allelic effects instead of having just one fixed effect per allele. The analyses showed that the estimated variance contributed by the QTL is sensitive to the choice of the priors. The susceptibility to prior information suggests that the data do not contain sufficient information to yield accurate estimates of the variance contributed by the QTL. More accurate estimates can be obtained by increasing the number of phenotypic observations. In order to obtain accurate estimates of the variance explained by the QTL, large designs are required. This is especially true for small QTL effects.

**Implications**

A new Bayesian method was developed that enables analysis of QTL accounting for heterogeneity of variance between sexes. Therefore, the method includes...
scale parameters for polygenic and QTL allelic effects per sex and residual variances per sex. In order to decrease computational requirements and improve estimability a parsimonious model specification is chosen and a reduced animal model is applied.

**Literature Cited**


**Appendix 1**

**Full Conditional Distributions of the Mean Effects in the Scaled Reduced Animal Model**

In scalar notation the distribution of the observations in the scaled RAM is as follows:

\[ y_{isj} | b, u, v, c, d, \sigma^2_{e_i} - N(X_i'b + c_iZ_u'u + d_iw_{ij}, \tau^2_{il}) \]

and in vector notation the distribution is:

\[ y | b, u, v, c, d, \sigma^2_{e_i} - N(Xb + Zu + Wv, T) \]

where \( \hat{Z} \) and \( \hat{W} \) are the matrices formed by concatenation of \( c_iZ_{isj} \) and \( d_iw_{isj} \), respectively. For example, depending on an animal's sex the nonzero elements of an animal's row of \( \hat{W} \) will be \( d_m \) or \( d_f \) for parents and \( 0.5d_m \) or \( 0.5d_f \) for nonparents. Finally, \( T \) is a diagonal matrix of the residual variances (including Mendelian sampling terms) corresponding to the observations (i.e., \( \tau^2_{ij} = \sigma^2_e + \omega(e^2 + 2d^2) \) for \( y_{isj} \)).

The following notations will be used: \( M_i \) denotes the \( i \)th column of matrix \( M \), \( M_{-,i} \) denotes matrix \( M \) with the \( i \)th column deleted, and \( m_{-,i} \) denotes vector \( m \) with the \( i \)th element deleted.

The full conditional distribution of the fixed effects is as follows:

\[ b_i | b_{-,i}, u, v, y, c, d, T - N(X'T^{-1}(y - X\hat{b} - \hat{Z}u - \hat{W}v), (X'T^{-1}X)^{-1}) \]
The full conditional distribution of the polygenic effects is as follows:

\[
    u_i \mid b, u_{-i}, v, y, c, d, T \sim N\left(\frac{\bar{Z}_i T^{-1}(y - \bar{X}_i b - \bar{Z}_i u_{-i} - \bar{W}_i v) - \sum_{j=1}^{4} a^u_j}{\bar{Z}_i T^{-1} \bar{Z}_i + a^u}, \frac{(\bar{Z}_i T^{-1} \bar{Z}_i + a^u)^{-1}}{\bar{Z}_i T^{-1}}\right)
\]

The full conditional distribution of the QTL allelic effects is as follows:

\[
    \nu_i \mid b, u, v_{-i}, y, c, d, T \sim N\left(\frac{\bar{W}_i T^{-1}(y - \bar{X}_i b - \bar{Z}_i u - \tilde{W}_i \nu_{-i}) - \sum_{j=1}^{4} g_j^u \nu_j}{\bar{W}_i T^{-1} \bar{W}_i + g_j^u}, \frac{(\bar{W}_i T^{-1} \bar{W}_i + g_j^u)^{-1}}{\bar{W}_i T^{-1}}\right)
\]

Appendix 2

Metropolis-Hastings Algorithm and Full Conditional Distributions of the Dispersion Parameters in the Scaled Reduced Animal Model

The conditional distributions of the dispersion parameters in the scaled RAM are as follows:

\[
p(c_s) = L_s \times (\sigma_{c_s}^2)^{-0.5} \exp\left(\frac{-0.5(c_s - \mu_{c_s})^2}{\sigma_{c_s}^2}\right) \text{ with } c_s \geq 0
\]

\[
p(d_s) = L_s \times (\sigma_{d_s}^2)^{-0.5} \exp\left(\frac{-0.5(d_s - \mu_{d_s})^2}{\sigma_{d_s}^2}\right) \text{ with } d_s \geq 0
\]

\[
p(\sigma_{e_s}^2) = L_s \times (\sigma_{e_s}^2)^{-\alpha_s-1} \exp\left(\frac{-\lambda_s}{\sigma_{e_s}^2}\right) \text{ with } \sigma_{e_s}^2 \geq 0
\]

The scalar Metropolis-Hastings algorithm applied to sample the dispersion parameters consists of 1) sampling a new candidate value using the candidate generating function, 2) calculating the values for the prior, 3) candidate generating density and 4) likelihood and, finally, 5) evaluating the Metropolis-Hastings ratio in order to decide whether the new sample is accepted or rejected. These five points are handled as follows:

1. For the scale parameters the candidate generating density was a normal distribution. For the error variances a uniform candidate generating density was applied.
2. For the scale parameters a normal prior distribution was applied. For the error variances an inverse Gamma prior distribution was applied.
3. For the scale parameters the candidate generating density of the new candidate vs the current value was assumed equal; therefore, the candidate generating density was dropped from the Metropolis-Hastings ratio. For the error variances a uniform candidate generating density was assumed. Therefore, the length of the interval in which the error variances are sampled is used as candidate generating density in the Metropolis-Hastings ratio.
4. The total likelihood of the RAM can be written as a multiplication of the likelihood per sex, as follows:

\[
    \prod_{s=m,f} L_s = \prod_{s=m,f} \prod_{i=1}^4 \left(\tau_{is}^{-0.5} \times \exp\left(-\frac{0.5 \sum_{j=1}^{n_{ij}} e_{isj}^2}{\tau_{is}}\right)\right)
\]

where \(e_{isj} = y_{isj} - x_{isj}'b - c_s z_{isj}^u - d_s u_{isj} \nu \) is the conditional residual of \(y_{isj}\) for the observed animal \(j\) with sex \(s\) in the RAM category \(i\) with residual variance \(\tau_{is} = \sigma_{e_s}^2 + \omega_1(e_{isj}^2 + 2d_s^2)\) and \(\omega_1\) reflects the total amount of additive genetic variance present in \(\tau_{is}\). There is one RAM category for parents (\(\omega_1 = 0\)) and three for nonparents: both parents known (\(\omega_2 = 0.5\)), one parent known (\(\omega_3 = 0.75\)), and both parents unknown (\(\omega_4 = 1.0\)). For parents, there are “ones” in \(z_{isj}\) and \(\omega_{isj}\) corresponding to the individuals’ own genetic effects, and for nonparents these are “halves” corresponding to the genetic effects of the identified parent(s) and “zeros” corresponding...
to unidentified parent(s). Note that if parents have no phenotypic observations, as in our case, the model reduces to a sire-dam model. The vectors $\mathbf{u}$ and $\mathbf{v}$ only contain parental genetic effects and the relationship matrices in their priors only contain parental contributions, hence $\mathbf{u} \sim N(0, \mathbf{A}_P)$ and $\mathbf{v}|m,k \sim N(0, \mathbf{G}_P)$. In the Metropolis-Hastings algorithm for scale parameters as well as error variances each time one parameter is sampled, therefore only the likelihood for one sex, $L_s$, needs to be evaluated. For each parameter, this likelihood needs to be evaluated twice with fixed $\mathbf{b}$, $\mathbf{u}$, and $\mathbf{v}$ in the Metropolis-Hastings updates. This is facilitated by computing the individual sums of squares and cross-products in the exponential term of each $L_s$, e.g.,

\[ \sum_j (y_{isj} - x_{isj}b)(z_{isj}u)^2 \text{ and } \sum_j (w_{isj}v)^2. \]

5. The Metropolis-Hastings ratio for the scale parameters and error variances is used to calculate the acceptance probability for the alternative:

\[
P(\text{acceptance}) \Rightarrow \begin{cases} 
\text{if } \frac{\text{prior}_{s,\text{current}} \times \text{cgd}_{s,\text{current}} \times L_{s,\text{current}}}{\text{prior}_{s,\text{alternative}} \times \text{cgd}_{s,\text{alternative}} \times L_{s,\text{alternative}}} > 0: \\
\frac{\text{prior}_{s,\text{alternative}} \times \text{cgd}_{s,\text{alternative}} \times L_{s,\text{alternative}}}{\text{prior}_{s,\text{current}} \times \text{cgd}_{s,\text{current}} \times L_{s,\text{current}}}, 1 \\
\text{otherwise:} \\
P(\text{acceptance}) = 1
\end{cases}
\]