Use of Multiple Genetic Markers in Prediction of Breeding Values

Johan A. M. van Arendonk,* Bruce Tier† and Brian P. Kinghorn‡

**Department of Animal Breeding, Wageningen Agricultural University, 6700 AH Wageningen, The Netherlands, †Animal Genetics and Breeding Unit, and ‡Department of Animal Science, University of New England, Armidale, New South Wales 2351, Australia*

Manuscript received July 26, 1993
Accepted for publication January 20, 1994

ABSTRACT

Genotypes at a marker locus give information on transmission of genes from parents to offspring and that information can be used in predicting the individuals' additive genetic value at a linked quantitative trait locus (MQTL). In this paper a recursive method is presented to build the gametic relationship matrix for an autosomal MQTL which requires knowledge on recombination rate between the marker locus and the MQTL linked to it. A method is also presented to obtain the inverse of the gametic relationship matrix. This information can be used in a mixed linear model for simultaneous evaluation of fixed effects, gametic effects at the MQTL and additive genetic effects due to quantitative trait loci unlinked to the marker locus (polygenes). An equivalent model can be written at the animal level using the numerator relationship matrix for the MQTL and a method for obtaining the inverse of this matrix is presented. Information on several unlinked marker loci, each of them linked to a different locus affecting the trait of interest, can be used by including an effect for each MQTL. The number of equations per animal in this case is $2m + 1$ where m is the number of MQTL. A method is presented to reduce the number of equations per animal to one by combining information on all MQTL and polygenes into one numerator relationship matrix. It is illustrated how the method can accommodate individuals with partial or no marker information. Numerical examples are given to illustrate the methods presented. Opportunities to use the presented model in constructing genetic maps are discussed.

BEST linear unbiased prediction (BLUP) methods are currently used for the prediction of breeding values of animals in a large number of countries and species. The prediction of an animal's breeding value using this method is based on phenotypes of the animal itself and/or those of its relatives. When only observations on the trait of interest are considered, the contribution of observations on relatives to an animal's breeding value depends on the additive genetic relationship, *i.e.*, the proportion of genes shared in common by descent, and the heritability of the trait. Mixed models used for prediction of breeding values require that the inverse of the numerator relationship matrix between animals (**A**) is known, and this matrix is generally very large. HENDERSON (1976) described a method to write the inverse of **A** directly from pedigree records and inbreeding coefficients. This has enabled the use of improved methods for estimation of genetic parameters and prediction of the breeding value of animals. Recently, the concept of the numerator relationship matrix has been extended to the gametic relationship matrix (**G**) where paternal and maternal gametes of an animal are considered separately (SMITH and ALLAIRE 1985; TIER and SÖLKNER 1993a). The gametic relationship matrix has been used for constructing the relationship matrix due to dominance effects (SCHAEFFER *et al.* 1989; SMITH and MÄKI-TANILLA 1990) and for the analysis of gametic imprinting effects (GIBSON *et al.* 1988; SCHAEFFER *et al.* 1989; TIER and SÖLKNER, 1993a). In building the relationship matrix or its inverse, at either the animal or gametic level, no knowledge on the actual contribution of a parent to its offspring is used. Instead use is made of WRIGHT's (1922) inbreeding coefficients and the coefficients of relationship between animals.

The detection of microsatellites and the use of polymerase chain reaction (SAIKI *et al.* 1988) make it possible to identify differences between individuals in genotype at many genomic sites. These sites are called marker loci (M), and their alleles are genetic markers. Marker loci are not likely to be quantitative trait loci (QTL) themselves, but they may be linked to QTL (SOLLER 1978). The use of information on markers is expected to accelerate genetic progress through increasing accuracy of selection, reduction of generation interval and increasing selection differentials (SOLLER and BECKMANN, 1983; KASHI *et al.* 1990; MEUWISSEN and VAN ARENDONK 1992).

FERNANDO and GROSSMAN (1989) showed how information on a single marker could be used in a mixed model analysis fitting additive effects for alleles at a QTL linked to the marker and additive effects for alleles at the remaining quantitative trait loci. They discussed the expansion of the model to include information on multiple markers but this expansion could not be applied to a large data set, unless a simple algorithm to invert the covariance matrix of all marked QTL is available. GODDARD (1992) extended the model of FERNANDO and

GROSSMAN (1989) to a situation with many linked markers and where genes for quantitative traits were bracketed between two markers. In an animal model with m quantitative trait loci, there are $2m+1$ effects to be estimated for each animal. The number of equations limits the application of the full model to relatively small data sets. The number of effects to be estimated can be greatly reduced by use of a reduced animal model (CANTET and SMITH 1991; GODDARD 1992). In this case effects are only predicted for animals that are parents. Breeding values and additive QTL effects for non-parents can be obtained by back solution when covariances between different QTL effects are zero.

TIER and SÖLKNER (1993) presented a general approach for constructing relationship matrix and its inverse based on partitioned matrix theory. In the present paper it will first be shown how this method can be used to incorporate information on a single autosomal marker locus in prediction of breeding values. In addition, a method will be presented to use information on multiple markers while predicting only one random effect for each individual.

METHOD

Inverse numerator relationship matrix without markers: The usual model to obtain BLUP of additive genetic effects for animal i (a_i), given phenotypic information, is

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + a_i + e_i \qquad (1)$$

where y_i is the phenotypic value of individual i, $\mathbf{x}_i'$ is a known incidence vector and $\boldsymbol{\beta}$ is a vector with fixed effects, and e_i is a random error. BLUP allows information from relatives to contribute to the predictor of a_i through the covariance matrix of a values. This covariance matrix is proportional to the numerator relationship matrix, **A**, which describes the genetic relationships between all individuals in the population. This matrix can be built recursively from a chronologically ordered list of pedigrees. A matrix representation of the rules for building $\mathbf{A}_i$ for animals 1 to i is

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{A}_{i-1} & \mathbf{A}_{i-1}\mathbf{s}_i \\ \mathbf{s}_i'\mathbf{A}_{i-1} & a_{ii} \end{bmatrix} \qquad (2)$$

where

$\mathbf{s}_i$ is a column vector with $i-1$ elements containing two elements $\frac{1}{2}$ corresponding to the sire and dam (if known) and zeroes elsewhere;

$\mathbf{A}_{i-1}$ is the numerator relationship matrix for animals 1 to $(i-1)$;

a_{ii} is the ith diagonal of **A** which is equal to $1 + F_i$ where F_i is the inbreeding coefficient of the ith animal.

TIER and SÖLKNER (1993) applied partitioned matrix theory to determine the effect of adding an additional row to **A** on the elements of $\mathbf{A}^{-1}$. They showed that

$$\mathbf{A}_i^{-1} = \begin{bmatrix} \mathbf{A}_{i-1}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} + (a_{ii} - \mathbf{s}_i'\mathbf{A}_{i-1}\mathbf{s}_i)^{-1} \begin{bmatrix} \mathbf{s}_i\mathbf{s}_i' & -\mathbf{s}_i \\ -\mathbf{s}_i' & 1 \end{bmatrix}. \qquad (3)$$

When both parents of i are known $\mathbf{s}_i'\mathbf{A}_{i-1}\mathbf{s}_i$ is equal to $\frac{1}{4}(a_{pp} + a_{pq} + a_{pq} + a_{qq})$, where a with subscripts are the four elements of $\mathbf{A}_{i-1}$ relating to the sire (p) and dam (q) of i. As a result $(a_{ii} - \mathbf{s}_i'\mathbf{A}_{i-1}\mathbf{s}_i)^{-1}$ can be rewritten as $(1 - \frac{1}{4}(a_{pp} + a_{qq}))^{-1}$. B. TIER and J. SÖLKNER (unpublished) showed that (3) is equivalent to the result of HENDERSON (1976) and QUAAS (1976). Equation 3, however, has the advantage that it is not restricted to the special form of $\mathbf{s}_i$ which contains only $\frac{1}{2}$ as non zero elements.

Inverse of gametic relationship matrix: Under an additive genetic model, the genetic value of the ith individual (a_i) is the sum of the additive genetic contribution from the paternally derived (g_i^p) and maternally (g_i^m) derived gametes, *i.e.*, $a_i = g_i^p + g_i^m$. The model for an observation on the ith individual can be written as

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + g_i^p + g_i^m + e_i. \qquad (4)$$

The variance-covariance matrix of **g** is proportional to the gametic relationship matrix (**G**) which can be set up using Equation 2 and building the equations by gametes. Elements of **G**, g_{ij}, describe the probability that alleles drawn from two gametes (i and j) are identical by descent. The diagonal elements of **G** are, therefore, all equal to one. We will illustrate construction of **G** and its inverse for the small population in Table 1 which is used throughout this paper. For animal 1, let g_1^p represent the paternally derived gamete and g_1^m represent the maternally derived gamete, and likewise for the other animals. The gametic relationships for the five animals are given in the lower half of Table 2. To construct $\mathbf{G}_{3^p}$, for example, the following vector is used $s_{3^p}' = [\frac{1}{2}, \frac{1}{2}, 0, 0]$. The matrix **A** may be obtained from **G** as

$$\mathbf{A} = \tfrac{1}{2}\mathbf{K}\mathbf{G}\mathbf{K}' \qquad (5)$$

where $\mathbf{K} = \mathbf{I}_n * [1, 1]$, where n is the number of individuals and $*$ denotes the Kronecker product of two matrices (SMITH and ALLAIRE 1985; TIER and SÖLKNER 1993a). The relationship between **a** and **g** is given by

$$\mathbf{a} = \mathbf{K}\mathbf{g}. \qquad (6)$$

TABLE 1

Example data for construction of $\mathbf{G}_v$ and $\mathbf{A}_v$ and their inverses

Animal	Sire[a]	Dam	Genotype M[b]	Phenotype
1	—	—	12	80
2	—	—	34	120
3	1	2	13	90
4	1	2	23	110
5	3	4	33	115

[a] Missing parent is indicated by —.

[b] Genotype at marker locus (M) is represented by combination alleles which are numbered from 1 to 4. In animals 1 and 2 the first marker allele is assigned to be linked to paternal QTL allele.

TABLE 2

Gametic relationship matrix for a single QTL using information on linked marker locus ($\mathbf{G}_{v|r}$, above diagonal) and without using marker information (G, below diagonal) for pedigree and data in Table 1 and recombination rate between marker and MQTL of 0.10

		1		2		3		4		5	
		p	*m*	*p*	*m*	*p*	*m*	*p*	*m*	*p*	*m*
1	*p*	1	0	0	0	0.9	0	0.1	0	0.09	0.01
	m	0	1	0	0	0.1	0	0.9	0	0.01	0.09
2	*p*	0	0	1	0	0	0.9	0	0.9	0.81	0.81
	m	0	0	0	1	0	0.1	0	0.1	0.09	0.09
3	*p*	0.5	0.5	0	0	1	0	0.18	0	0.1	0.02
	m	0	0	0.5	0.5	0	1	0	0.82	0.9	0.74
4	*p*	0.5	0.5	0	0	0.5	0	1	0	0.02	0.1
	m	0	0	0.5	0.5	0	0.5	0	1	0.74	0.9
5	*p*	0.25	0.25	0.25	0.25	0.5	0.5	0.25	0.25	1	0.67
	m	0.25	0.25	0.25	0.25	0.25	0.25	0.5	0.5	0.25	1

Diagonal elements $\mathbf{G}_v$ and $\mathbf{G}$ are identical and given on diagonal.

There is, however, no closed form relationship between $\mathbf{A}^{-1}$ and $\mathbf{G}^{-1}$ nor is it possible to form $\mathbf{G}$ from $\mathbf{A}$, directly.

Gametic relationship matrix for a single QTL linked to a marker locus: Consider a single polymorphic marker locus (M), closely linked to a quantitative trait locus (QTL). This QTL will be referred to as a marked QTL (MQTL) to distinguish it from any other QTL affecting the trait. The information on M gives information on transmission of genes from parents to offspring and that information can be used in predicting the individuals' additive gametic effects at the linked QTL. The model for an observation on individual i can be written as

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + v_i^p + v_i^m + u_i + e_i \qquad (7)$$

where v_i^p and v_i^m represent the additive gametic effects at the MQTL and u_i the additive genetic effects due to QTL unlinked to M. The covariance matrix of v_i values with marker information is proportional to the gametic relationship matrix for the MQTL ($\mathbf{G}_{v|r}$) which depends on the recombination rate (r) between M and MQTL and on marker genotype information on individuals. Without marker information, v_i^p has a relationship of $\frac{1}{2}$ with both the paternal and the maternal gametic effect of its sire. Information on M which is linked to the MQTL changes these relationships when transmission of marker allele from sire to offspring can be traced. This will first be illustrated for some small pedigrees given in Table 3. The notation used to identify alleles is given in Figure 1. Offspring i has one marker allele from the sire (m_i^p) and one from the dam (m_i^m). From the genotypic information on the M, probabilities of each marker allele coming from the sire and the dam are calculated first. Secondly the probabilities that the m_i^p is identical by descent to m_s^p and m_s^m are determined. In the first pedigree given in Table 3, marker allele 1 in the offspring ($m_1 = 1$) comes from the sire, *i.e.* $P(m_1 \equiv m_s) = 1$, and m_1 is identical by descent to the allele in the sire which is assigned to be m_s^p. The second marker allele is identical by descent to m_d^p.

Given the information on the marker genotypes and the recombination rate between M and QTL (r) we can determine the relationships of v_i^p with v_s^p and v_s^m. For pedigree 1 in Table 3 the relationships for v_i^p are $p(v_i^p \equiv v_s^p) = (1 - r)$ and $p(v_i^p \equiv v_s^m) = r$, while for v_i^m they are $p(v_i^m \equiv v_d^p) = (1 - r)$ and $p(v_i^m \equiv v_d^m) = r$. These probabilities can be used in $\mathbf{s}$ for building the gametic relationship matrix ($\mathbf{G}_{v|r}$) and its inverse ($\mathbf{G}_{v|r}^{-1}$) for the marked QTL.

Pedigree 6 in Table 3 contains only information on the allele transmitted by the dam but not by the sire. The vector $\mathbf{s}'$ for the paternal gamete of the offspring in that case is $[\frac{1}{2}, \frac{1}{2}, 0, 0]$, while for the maternal gamete $\mathbf{s}' = [0, 0, (1 - r), r, 0]$.

When gametes are ordered chronologically, the gametic relationship matrix for gametes 1 to k ($\mathbf{G}_{v|r,k}$) can be obtained from

$$\mathbf{G}_{v|r,k} = \begin{bmatrix} \mathbf{G}_{v|r,k-1} & \mathbf{G}_{v|r,k-1}\mathbf{s}_k \\ \mathbf{s}_k'\mathbf{G}_{v|r,k-1} & g_{kk} \end{bmatrix} \qquad (8)$$

where

- $\mathbf{s}_k$ is column vector ($k - 1$ elements) containing non-zero elements relating gamete k to gametes in the parent (if known) and zeroes elsewhere;
- $\mathbf{G}_{v|r,k-1}$ is gametic relationship matrix for gametes 1 to $(k - 1)$;
- g_{kk} is the diagonal of $\mathbf{G}_v$ relating to the kth gamete which is equal to one.

The inverse of $\mathbf{G}_{v|r,k}$ can be obtained from

$$\mathbf{G}_{v|r,k}^{-1} = \begin{bmatrix} \mathbf{G}_{v|r,k-1}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} + (g_{kk} - \mathbf{s}_k'\mathbf{G}_{v|r,k-1}\mathbf{s}_k)^{-1} \begin{bmatrix} \mathbf{s}_k\mathbf{s}_k' & -\mathbf{s}_k \\ -\mathbf{s}_k' & 1 \end{bmatrix}. \qquad (9)$$

TABLE 3

Probability of marker alleles in offspring (m_1, m_2) being equal to parental marker alleles in sire (m_s^p, m_s^m) and dam (m_d^p, m_d^m) for several pedigrees (Ped.) differing in genotype of sire, dam and/or offspring

					Prob(m_1)[d]				Prob(m_2)[d]			
Ped.	Sire[a]	Dam[a]	Off[a,b]	$P(m_1 \equiv m_s)$[c]	m_s^p	m_s^m	m_d^p	m_d^m	m_s^p	m_s^m	m_d^p	m_d^m
1	12	34	13	1	1	0	0	0	0	0	1	0
2	12	12	11	½	½	0	½	0	½	0	½	0
3	12	12	12	½	½	0	½	0	0	½	0	½
4	11	22	12	1	½	½	0	0	0	0	½	½
5	11	12	12	1	½	½	0	0	0	0	0	1
6	11	23	12	1	½	½	0	0	0	0	1	0

[a] Genotype is represented by combination of alleles which are numbered from 1 to 4. For sire and dam the first allele is assigned to be the paternal marker allele, *i.e.*, m_s^p and m_d^p, respectively.
[b] The first allele in offspring is referred to as m_1 and the second as m_2.
[c] Probability that m_1 is identical by descent to an allele in sire.
[d] Probability that allele in offspring is identical by descent to the four parental alleles.

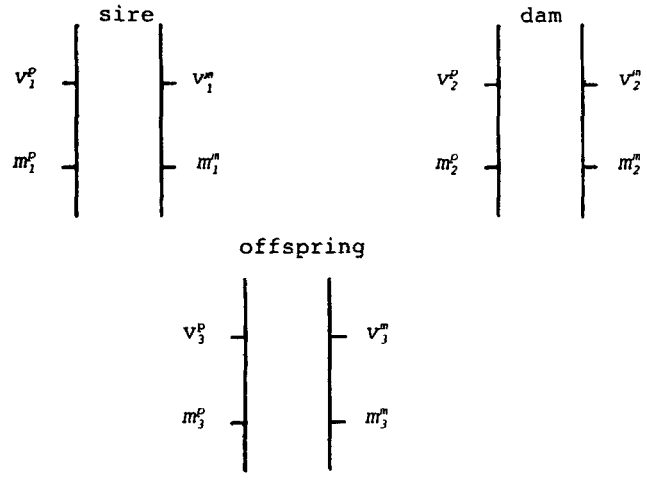


FIGURE 1.—Pedigree with sire (1), dam (2) and offspring (3) where v are alleles at quantitative trait locus and m are alleles at marker locus.

Numerator relationship matrix using a single marker: Observations are available on animals and not on gametes but an animal level model can be written which is equivalent to the gametic level model. To be able to use a model at the animal level, the inverse of the numerator relationship matrix for MQTL is needed. Similarly to (5), the numerator relationship matrix for the MQTL for animals 1 to i ($\mathbf{A}_{v|r,i}$) can be obtained from the gametic relationship matrix $\mathbf{G}_{v|r,2i}$ using

$$\mathbf{A}_{v|r,i} = \tfrac{1}{2}\mathbf{K}\mathbf{G}_{v|r,2i}\mathbf{K}' \qquad (10)$$

where $\mathbf{K} = \mathbf{I}_i * [1, 1]$, i is the number of individuals and * denotes the Kronecker product of two matrices. To enable the use of Equation 3 for obtaining $\mathbf{A}^{-1}_{v|r,i}$ the vector $\mathbf{s}_i$ reflecting the ancestral contributions is needed. From Equation 2 it can be seen that the row vector containing the numerator relationships of i with animals 1 to $i-1$ ($\mathbf{A}^{i}_{v|r,i}$) is equal to

$$\mathbf{A}^{i}_{v|r,i} = \mathbf{A}_{v|r,i-1}\mathbf{s}_i. \qquad (11)$$

Consequently, the vector with ancestral contributions can be obtained from

$$\mathbf{s}_i = \mathbf{A}^{-1}_{v|r,i-1}\mathbf{A}^{i}_{v|r,i}. \qquad (12)$$

Given $\mathbf{s}_i$ we can use Equation 3 to get $\mathbf{A}^{-1}_{v|r,i}$.

Estimation of breeding values: In matrix notation model (7) can be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{v} + \mathbf{e} \qquad (13)$$

where

- $\mathbf{y}$ is the vector of observations on the trait of interest,
- $\boldsymbol{\beta}$ is the vector with fixed effects,
- $\mathbf{u}$ is the random vector with additive genetic effects due to loci not linked to the M,
- $\mathbf{v}$ is the random vector with gametic effects at the MQTL,
- $\mathbf{e}$ is the random vector of residual effects.

The matrices $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{W}$ are incidence matrices and the variance-covariance structure of the random variables is

$$V\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_u\sigma_u^2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_{v|r}\sigma_v^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2\end{bmatrix} \qquad (14)$$

where $\mathbf{A}_u$ is the numerator relationship matrix for the QTL which are unlinked to M, $\mathbf{G}_{v|r}$ is the gametic relationship matrix for the MQTL, and $\mathbf{I}$ is an identity matrix. Also, σ_u^2, σ_v^2 and σ_e^2 are the additive genetic variance due to QTL not linked to M, the variance due to gametic effects at the MQTL and the residual variance, respectively. The total additive genetic variance (σ_a^2) is equal to $\sigma_a^2 = \sigma_u^2 + 2\sigma_v^2$. Let $\alpha_u = \sigma_e^2/\sigma_u^2$ and $\alpha_v = \sigma_e^2/\sigma_v^2$, then the mixed model equations of (13) are

$$\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W}\\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}_u^{-1}\alpha_u & \mathbf{Z}'\mathbf{W}\\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{Z} & \mathbf{W}'\mathbf{W} + \mathbf{G}^{-1}_{v|r}\alpha_v\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{u}}\\ \hat{\mathbf{v}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{y}\\ \mathbf{Z}'\mathbf{y}\\ \mathbf{W}'\mathbf{y}\end{bmatrix}. \qquad (15)$$

For selection of animals it is relevant to know the predicted

total additive genetic merit, *i.e.*, the sum of the predicted value for polygenic effects and predicted gametic effects. Predicted total additive genetic merit ($\hat{\mathbf{a}}$) can be calculated as

$$\hat{\mathbf{a}} = \hat{\mathbf{u}} + \mathbf{K}\hat{\mathbf{v}}. \tag{16}$$

The matrix $\mathbf{K}$, which relates gametic effects to animals, is identical to $\mathbf{W}$ in (13) when all individuals have observations.

The variance covariance matrix of $\mathbf{a}$ calculated from $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{V}_{a|r}$ which uses information on r, is equal to

$$\begin{aligned}\mathbf{V}_{a|r} &= \text{Var}(\mathbf{u} + \mathbf{K}\mathbf{v}|\, r)\\ &= \text{Var}(\mathbf{u}|\, r) + \text{Var}(\mathbf{K}\mathbf{v}|\, r)\\ &= \text{Var}(\mathbf{u}) + \mathbf{K}\,\text{Var}(\mathbf{v}|\, r)\mathbf{K}'\\ &= \mathbf{A}_u\sigma_u^2 + \mathbf{K}\mathbf{G}_{v|r}\mathbf{K}'\sigma_v^2\\ &= \mathbf{A}_u\sigma_u^2 + 2\mathbf{A}_{v|r}\sigma_v^2.\end{aligned}$$

The variance due to additive gametic effects at the MQTL (σ_v^2) is half the additive genetic variance due to the MQTL (σ_q^2) and, as a result

$$\mathbf{V}_{a|r} = \mathbf{A}_u\sigma_u^2 + \mathbf{A}_{v|r}\sigma_q^2. \tag{17}$$

The combined numerator relationship matrix including information on the MQTL ($\mathbf{A}_{a|r}$) is equal to

$$\mathbf{A}_{a|r} = \mathbf{V}_{a|r}\sigma_a^{-2} = \mathbf{A}_u\sigma_u^2/\sigma_a^2 + \mathbf{A}_{v|r}\sigma_q^2/\sigma_a^2. \tag{18}$$

The combined additive genetic merit ($\mathbf{a}$) can be predicted directly from the following model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}. \tag{19}$$

The mixed model equations of (19) are

$$\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}\\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}_{a|r}^{-1}\alpha_a\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{a}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{y}\\ \mathbf{Z}'\mathbf{y}\end{bmatrix} \tag{20}$$

where $\alpha_a = \sigma_e^2/\sigma_a^2$. The inverse $\mathbf{A}_{a|r}^{-1}$ can be obtained using Equations 12 and 3.

Several unlinked markers: If information on several (say, m) MQTL is available, model (7) can be extended to include gametic effects for each MQTL

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \sum_{k=1}^{m}(v_{ik}^p + v_{ik}^m) + u_i + e_i. \tag{21}$$

This approach, however, results in two additional equations for each individual for each MQTL introduced in the analysis. This results in $2m + 1$ equation per individual, where m is the number of MQTL. For a large number of individuals and a large number of MQTL, solving the mixed model equations may not be feasible (FERNANDO and GROSSMAN 1989). Instead of fitting gametic effects for each MQTL, the additive genetic effect for each MQTL might be fitted, in which case the number of equations reduces to $m + 1$. The number of equations is reduced to 1 per animal, however, when just the sum of additive effects of all QTL (a_i) is included in the model

$$a_i = \sum_{k=1}^{m}(v_{ik}^p + v_{ik}^m) + u_i. \tag{22}$$

The numerator relationship matrix for total additive genetic merit can be obtained using

$$\mathbf{A}_{a|rm} = \mathbf{A}_u\frac{\sigma_u^2}{\sigma_a^2} + \sum_{k=1}^{m}\mathbf{A}_{v|rk}\frac{\sigma_{qk}^2}{\sigma_a^2} \tag{23}$$

where $\mathbf{A}_{v|rk}$ is the numerator relationship matrix for the kth MQTL and σ_{qk}^2 is the additive genetic variation explained by the kth MQTL. The recombination rate for building $\mathbf{A}_{v|rk}$ will differ between MQTL. Equations 7 to 12 can be used to construct $\mathbf{A}_{v|rk}$. The inverse of $\mathbf{A}_{a|rm}$ can be obtained using Equations 12 and 3.

NUMERICAL EXAMPLES

Gametic relationship matrices and their inverse: Consider the pedigree and genotypes for M in Table 1. To construct $\mathbf{G}_{v|r}$ and $\mathbf{A}_{v|r}$ we take $r = 0.1$. The gametic relationship matrix for the single QTL using information on linked M is given in Table 2. Animal 3 inherited marker allele m_s^p from the sire. As a consequence, the paternally derived gametic effect at the QTL of animal 3 (v_3^p) has a relation of $(1 - r)$ with v_1^p and r with v_1^m. Animal 3 and 4 are full sibs which inherited different marker alleles from the sire but the same marker allele from the dam. For these full sibs to inherit the same paternal allele at the QTL, recombination needs to have occurred in the formation of the gamete transmitted to one offspring and no recombination occurred in that transmitted to the other, which has a probability of $2 * (1 - r) * r = 0.18$ (Table 2). The QTL allele transmitted by the dam to animals 3 and 4 is identical when no recombination or recombination occurred in both gametes produced, of which the probability is $(1 - r) * (1 - r) + r * r = 0.82$. It can be seen from this example that the relationship between gametic effects in full sibs can deviate from $\frac{1}{2}$, which is the value found with no information on a linked marker locus.

Animal 5 is an offspring from full sibs 3 and 4. With no marker information the probability that the maternal and paternal alleles at the QTL are identical by descent is $\frac{1}{4}$ (Table 2). Animal 5 inherited marker allele 3 from both parents which results in a relationship of 0.666 which can be explained as follows. The parents received allele 3 from dam 2. Both v_5^p and v_5^m are identical by descent to v_2^p when no recombination has occurred when genes where transmitted from 2 to 3 and 4, and the subsequent transmission from 3 and 4 to 5: $P(v_5^p \equiv v_2^p$ and $v_5^m \equiv v_2^p) = 0.9^4 = 0.6561$. Secondly, alleles can be identical by descent when recombination occurred in gametes transmitted by 2 to 3 and 4 but no recombination thereafter, *i.e.* $P(v_5^p \equiv v_2^m$ and $v_5^m \equiv v_2^m) = 0.1^2 * 0.9^2 =$

TABLE 4

Inverse of gametic relationship matrix for a single QTL using information on linked marker locus ($\mathbf{G}^{-1}_{v|r}$, off diagonal elements in upper triangle) and without using marker information ($\mathbf{G}^{-1}$, off diagonal elements in lower triangle) for pedigree and marker information in Table 1 and recombination rate between marker and MQTL of 0.10

		Diagonal		1		2		3		4		5	
		$\mathbf{G}^{-1}$	$\mathbf{G}^{-1}_{v\|r}$	p	m	p	m	p	m	p	m	p	m
1	p	2	5.56		1	0	0	−5	0	−0.56	0	0	0
	m	2	5.56	1		0	0	−0.56	0	−5	0	0	0
2	p	2	10.00	0	0		1	0	−5	0	−5	0	0
	m	2	1.11	0	0	1		0	−0.56	0	−0.56	0	0
3	p	2.5	5.61	−1	−1	0	0		0.50	0	0	−0.56	0
	m	2.5	10.06	0	0	−1	−1	0.5		0	0	−5	0
4	p	2.5	5.61	−1	−1	0	0	0	0		0.50	0	−0.56
	m	2.5	10.06	0	0	−1	−1	0	0	0.5		0	−5
5	p	2	5.56	0	0	0	0	−1	−1	0	0		0
	m	2	5.56	0	0	0	0	0	0	−1	−1	0	

0.0018. In addition, v_5^p can be identical by descent to v_5^m due to alleles transmitted by sire 1, but this probability is small given that different marker alleles where transmitted to 3 and 4: $P(v_5^p \equiv v_1^p \text{ and } v_5^m \equiv v_1^p) = 0.9 * 0.1 * 0.1 * 0.1 = 0.0008$ and $P(v_5^p \equiv v_1^m \text{ and } v_5^m \equiv v_1^m) = 0.1 * 0.9 * 0.1 * 0.1 = 0.0008$. In this case, recombination is needed in one of the gametes transmitted by sire 1 and in both gametes transmitted by 3 and 4. All four possibilities result in $P(v_5^p \equiv v_5^m) = 0.6561 + 0.0081 + 0.0009 + 0.0009 = 0.666$ which is equal to the value given in Table 2.

The inverse of the gametic relationship matrix for a single MQTL using information on linked M ($\mathbf{G}^{-1}_{v|r}$) and without using marker information, the latter being equal to $\mathbf{G}^{-1}$, are given in Table 4. Non-zero elements in $\mathbf{G}^{-1}_{v|r}$ and $\mathbf{G}^{-1}$ occur at the same positions but their values are different. In $\mathbf{G}^{-1}$ the two elements relating a gametic effect to its parent are equal to -1. In $\mathbf{G}^{-1}_{v|r}$ these elements are equal to -0.56 and -5.56. In $\mathbf{G}^{-1}$ the diagonal elements for animals 1 and 2 are identical which reflects that the same amount of information is available to estimate the gametic effects. As a result of using marker information the amount of information available to estimate gametic effects in animals 1 and 2 differs and as a consequence the corresponding diagonal elements in $\mathbf{G}^{-1}_{v|r}$ are different. Dam 2 transmitted marker allele 3 to both offspring and as a consequence most information is available to estimate the paternal gametic effect in animal 2 which is associated with it (v_2^p), which is reflected by the large diagonal element of 10.

Numerator relationship matrices: Equation 10 can be used to calculate the numerator relationship matrix ($\mathbf{A}_{v|r}$) from the gametic relationship matrix for the QTL linked to the marker ($\mathbf{G}_{v|r}$). The results are given in Table 5 which also gives the numerator relationship matrix without marker information ($\mathbf{A}$). The relationships between animals 1 to 4 are identical with and without marker information. However, this is not the case in general. For example, the numerator relationship between full sib offspring of 1 and 2 with marker genotype 13 and 24 is only 0.18. The relationship between these full sibs at M is zero. Due to recombination between M and QTL, alleles at the QTL have probability of 0.18 of being identical by descent. When two full sib offspring of 1 and 2 have same marker genotypes (*e.g.*, 13) the relationship at M is 1 and at QTL 0.82. As expected, the numerator relationship of an offspring with its parents remains $\frac{1}{2}$ in all cases. For the marker genotypes given in Table 1, the relationship at M is $\frac{1}{2}$ and consequently the relationship at QTL is also $\frac{1}{2}$.

The relationship of animal 5 with animals 1 to 4 are changed to the extent that grandparent 2 has a much higher relation than grandparent 1 (Table 5). In addition, the diagonal element for animal 5 is increased. The animal is completely inbred at the marker locus and the inbreeding coefficient at the linked QTL is 0.666, which is by definition equal to $P(v_5^p \equiv v_5^m)$ in Table 2.

To obtain the inverse of $\mathbf{A}_{v|r}$ using (3), the vectors $\mathbf{s}_i$ need to be calculated first using (12). The vector $\mathbf{s}_i$ is given in Table 6 for animals (i) 2 to 5. For animals 3 and 4, $\mathbf{s}_i$ contained only non-zero elements of $\frac{1}{2}$ for their parents which makes these vectors identical to those for building $\mathbf{A}^{-1}$. This was to be expected because numerator relationships between 1 to 4 were identical in $\mathbf{A}_{v|r}$ and $\mathbf{A}$. The vector $\mathbf{s}_5$, however, contains non-zero elements for both parents and grandparents. The contribution of both parents is equal but the granddam (2) has a positive contribution while the contribution of the grandsire (1) is negative. The negative contribution reflects that the numerator relationship between 1 and 5 based on marker information is much smaller than expected based on the relationship ($\frac{1}{2}$) between 1 and the parents of 5 (3 and 4). The inverse of $\mathbf{A}_{v|r}$ and $\mathbf{A}$ are given in Table 7.

Prediction of effects: Let us look at the prediction of random effects for model (7) which includes one ran-

TABLE 5

Numerator relationship matrix for animals in Table 1 for a single MQTL using information on linked marker locus ($\mathbf{A}_{v|r}$) and without using marker information (A) for recombination rate (r) of 0.10

	With marker information ($\mathbf{A}_{v\|r}$)					No marker information (A)				
i	1	2	3	4	5	1	2	3	4	5
1	1	0	0.5	0.5	0.1	1	0	0.5	0.5	0.5
2	0	1	0.5	0.5	0.9	0	1	0.5	0.5	0.5
3	0.5	0.5	1	0.5	0.88	0.5	0.5	1	0.5	0.75
4	0.5	0.5	0.5	1	0.88	0.5	0.5	0.5	1	0.75
5	0.1	0.9	0.88	0.88	1.67	0.5	0.5	0.75	0.75	1.25

TABLE 6

Vector ($\mathbf{s}_i$) with contributions at the QTL linked to M from ancestors for animals (i) 2 to 5 in the pedigree given in Table 1 (r = 0.1)

Animal (i)	Elements in $\mathbf{s}_i$ relating to animal			
	1	2	3	4
2	0			
3	½	½		
4	½	½	0	
5	−0.656	0.144	0.756	0.756

dom effect due to polygenes unlinked to M (u_i) and two random effects for additive gametic effects at one MQTL. The following parameters (all expressed as proportion of phenotypic variation) were used: σ_u^2 = 0.3, σ_v^2 = 0.05, σ_a^2 = 0.3 + 2 * 0.05 = 0.40, σ_e^2 = 0.6, *i.e.*, 40% of phenotypic variation is due to additive genetic variation of which 25% can be explained by the MQTL.

In Table 8, predictions of combined additive genetic effect (a) are given with and without using information on M. When no marker information used, a value of 1.5 (σ_e^2/σ_a^2) was taken for α_a and the inverse of the numerator relationship matrix ($\mathbf{A}^{-1}$) is used in (20).

Predictions of combined additive genetic merit (a_i) could be obtained by solving equation (15) and summing the relevant estimates ($a_i = v_i^p + v_i^m + u_i$). Alternatively, Equation 20 could be used in which the combined numerator relationship matrix ($\mathbf{A}_{a|rm}$) is used which from (23) is equal to

$$A_{a|rm} = \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0.4 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0.6 \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{2} & 0.782 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0.782 \\ 0.4 & 0.6 & 0.782 & 0.782 & 1.354 \end{bmatrix}.$$

For animal 5, a increased considerably after using marker information as a result of the increased relationship with granddam 2 which had a high phenotype (Table 8).

DISCUSSION

TIER and SÖLKNER (1993) presented a method based on partitioned matrix theory to obtain the inverse of the numerator relationship matrix. In this paper it is shown how their method can be used to incorporate information on a marker locus (M) into building a relationship matrix for a QTL linked to a marker locus (MQTL). We have only considered the situation where all animals have genotypic information on one M. This situation is, however, not likely to occur in livestock. To illustrate how the procedure can be extended to allow for animals without marker information, a situation with a single marker locus will be considered. The Equations 8 and 9 can still be used to calculate the gametic relationship matrix for the MQTL and its inverse. The problem is to find the appropriate vectors **s**. With no marker genotype information on the offspring, **s** for both gametic effects in that offspring contains non-zero elements equal to $\frac{1}{2}$ relating the gametic effect in the offspring to those in the relevant parent. This is true independent of whether or not marker genotypes on the parents are available. In a situation where all animals are genotyped, marker alleles can be assigned to gametic effects in a base animal. This is not true for a genotyped animal which is an offspring of two ungenotyped parents. In that case each marker allele should have a probability of $\frac{1}{2}$ of being linked to either the paternal or maternal gametic effect in that animal. By doing this, gametic effects in subsequent genotyped offspring have equal covariances with gametic effects in their ungenotyped grandparents. The vector **s** can be derived by the procedure described in APPENDIX A.

The situation which needs special consideration is that where a genotyped offspring has one or both parents with unknown marker genotype and when one or both parents have more than one genotyped offspring. When, for example, the marker genotypes are not available on the dam, the covariance between maternal gametic effects in two genotyped offspring i and j can be derived as follows

$$g_{ij} = p(v_i^m \equiv v_j^m) \qquad (24)$$
$$= p(m_i^m \equiv m_j^m)\{(1 - r)^2 + r^2\} + p(m_i^m \not\equiv m_j^m)2(1 - r)r$$

TABLE 7

Inverse of numerator relationship matrix for animals in Table 1 for a single MQTL using information on linked M ($\mathbf{A}_{v|r}^{-1}$) and without using marker information ($\mathbf{A}^{-1}$) for recombination rate between marker 1 and QTL of 0.10

	With marker information ($\mathbf{A}_{v|r}^{-1}$)					No marker information ($\mathbf{A}^{-1}$)				
	1	2	3	4	5	1	2	3	4	5
1	3.57	0.66	−2.81	−2.81	2.39	2	1	−1	−1	0
2	0.66	2.08	−0.60	−0.60	−0.52	1	2	−1	−1	0
3	−2.81	−0.60	4.08	2.08	−2.75	−1	−1	2.5	0.5	−1
4	−2.81	−0.60	2.08	4.08	−2.75	−1	−1	0.5	2.5	−1
5	2.39	−0.52	−2.75	−2.75	3.64	0	0	−1	−1	2

TABLE 8

Prediction of effects for animals without and with using information on marker locus associated with MQTL

	Without marker	With marker information			
i	a_i	v_i^p	v_i^m	u_i	a_i
1	−8.00	−1.74	−0.74	−5.84	−8.32
2	8.00	1.57	0.91	5.84	8.32
3	−1.84	−1.78	1.46	−1.46	−1.78
4	3.16	−0.75	1.69	2.29	3.23
5	3.75	1.26	1.57	2.43	5.26

Pedigree and data from Table 1; $r = 0.1$, $\sigma_u^2 = 0.3$, $\sigma_q^2 = 0.1$, $\sigma_e^2 = 0.6$.

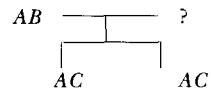
where $p(m_i^m \equiv m_j^m)$ is the probability that the maternally inherited marker alleles in i and j are identical by descent and $p(m_i^m \not\equiv m_j^m)$ is the probability that they are not equal by descent. The latter probabilities can be derived from the genotype of the sire, genotypes of i and j and population allelic frequencies. This is illustrated in Table 9. The row and column in the gametic relationship matrix for v_i^m can be built with a vector $\mathbf{s}$ with two elements equal to $\frac{1}{2}$ corresponding to the gametic effects in the dam. As shown in APPENDIX A, $\mathbf{s}$ for the maternal gametic effect in the second offspring (v_j^m) contains $(1 - g_{ij})$ at the elements corresponding to the gametic effects in the dam while there is also an element equal to $(2g_{ij} - 1)$ at the position of maternal gametic effect in i (v_i^m). In Table 9, it is illustrated how genotypic probabilities are used to calculate probability that marker alleles are identical by descent. For a situation with more than two offspring the gametic relationships can be determined using Equation 24 where genotypic probabilities of the parent(s) without information on M is used to determine $p(m_i^m \equiv m_j^m)$. Methods are available to efficiently determine genotypic probabilities for each M for animals with unobserved genotypes from information on pedigree and observed marker genotypes (VAN ARENDONK *et al.* 1989; FERNANDO *et al.* 1993; JANSS *et al.* 1993). With more than two offspring, the vector $\mathbf{s}$, which is needed for building the inverse using (9), can be obtained using the equivalent of (11) and (12) for gametic effects (APPENDIX A). The presented procedure can be extended to a situation with multiple markers as well as the combined matrix approach. It can, therefore, be concluded that the presented procedure is sufficiently general to incorporate individuals without information on marker loci.

LANDER and BOTSTEIN (1989) have pointed out that marker brackets are more efficient for estimating MQTL effects compared to an analysis based on either of the bracketing markers alone. With marker brackets, the probabilities that MQTL effects in different animals are identical by descent can be calculated more accurately. The procedure presented here can use information on two marker loci bracketing the MQTL or even on several markers linked to a single QTL. Consider two marker loci bracketing the MQTL and recombination rates between marker and MQTL of r_1 and r_2. The non-zero elements of $\mathbf{s}_k$ for the paternal gametic effect in an offspring which inherited the paternal marker bracket of the sire, is equal to $(1 - r_1)(1 - r_2) + r_1 r_2$ and $(1 - r_1) r_2 + (1 - r_2) r_1$ for the paternal and maternal gametic effect in the sire, respectively. Without complete knowledge of the linkage phase of marker alleles in the parent, all alternatives need to be weighted with their probability of occurrence given the available information. In that case a procedure similar to that in APPENDIX A can be used to determine the elements of $\mathbf{s}$.

The method presented in this paper is very similar to that of FERNANDO and GROSSMAN (1989) which cannot be applied in situations where parents are inbred and which requires assigning paternal and maternal origin of marker alleles. WANG *et al.* (1991) described a method which does not require assigning the origin of the alleles and accounts for inbreeding. Also the procedure presented here to obtain $\mathbf{G}_{v|r}$ and $\mathbf{G}_{v|r}^{-1}$ can be applied when parents are inbred. In addition, the method allows for individuals without information on marker locus. Moreover, the procedure presented by FERNANDO and GROSSMAN (1989) cannot be used in situations where the effects of more than one MQTL are combined and, therefore, require fitting two gametic effects for each animal in the mixed model equations. The current procedure can be applied to a situation with more than one MQTL. In addition, the number of random effects predicted for each animal can be reduced to one, *i.e.* the total additive genetic merit ($\mathbf{a}$). In obtaining the numerator relationship matrix $\mathbf{A}_{a|rm}$ using Equation 23, it is assumed that all MQTL are unlinked. In case of linkage between two

TABLE 9

Calculation of probabilities of identity by descent of maternally inherited marker alleles in two full sibs, $p(m_i^m = m_j^m)$, for a given pedigree where the dam is not genotyped for *M*



Dam genotype[a]	Probabilities			$p(m_i^m \equiv m_j^m)$	$p(m_i^m \not\equiv m_j^m)$
	Prior[b]	Joint[c]	Rescaled		
AB	¼	¼ * ¼²	⅙	1	0
BC	¼	¼ * ¼²	⅙	1	0
CC	¼	¼ * ½²	⅔	½	½
Weighted[d]				0.67	0.33

In the population there are three alleles for *M* (A, B and C) with a frequency of ¼, ¼ and ½.

[a] Only genotypes which based on offspring genotypes have a probability larger than zero are shown.

[b] Prior probabilities are calculated assuming a population in Hardy-Weinberg equilibrium.

[c] The probability of offspring having genotype *AC* given the parental genotypes *AB* and *AC* is equal to ¼. Combining the prior probability (¼) with the information on both offspring results in a joint probability of ¼ * ¼².

[d] The weighted probability $p(m_i^m \equiv m_j^m) = (1 + 1 + 4 * ½)/6 = 0.667$.

or more MQTL, Equation 23 needs to be expanded with terms relating to the covariances between effects at linked MQTL.

The method described by GODDARD (1992) is very similar to that of FERNANDO and GROSSMAN (1989) but allows for multiple marker brackets. In constructing the gametic relationship for a MQTL, it is assumed that no double recombinants occur between markers. It is not clear how uncertainty about marker haplotypes in parents or animals without marker information can be accounted for.

Application of the presented procedures requires knowledge of the recombination rate (r) between *M* and MQTL and the additive genetic variance explained by MQTL (σ_v^2 or σ_q^2). The models presented here can be used to estimate these parameters using derivative-free REML procedures (GRASER *et al.* 1987). This is not restricted to the effect of one QTL alone but allows for simultaneous estimation of parameters (r, σ_q^2) for different MQTL and unlinked QTL (σ_u^2) as described in more detail by VAN ARENDONK *et al.* (1993).

The costs of building the inverse of the combined numerator relationship matrix ($\mathbf{A}_{a|rm}^{-1}$) for very large populations (hundreds of thousands of animals) would be prohibitively expensive with current computer facilities. However, it is unlikely that marker information will be collected on such large numbers of animals in the near future. Given the current limitations for population size (thousands of animals) and number of random effects in the models for popular REML algorithms for estimating components of variance, the use of a combined numerator relationship matrix for some of the random effects should be tractable.

For large populations and numbers of MQTL the question arises whether it is more efficient to use the combined $\mathbf{A}_{a|rm}$ and hence one equation per animal or to use m $\mathbf{G}_{v|r}$ matrices and one $\mathbf{A}_u$ matrix and have $2m + 1$ equations per animal. Let there be m MQTL, f fixed effects, n animals in the complete model (21) and marker genotypes on all animals. Now consider the number of contributions from each observed animal to the half-stored mixed model equations. An additional animal with identified parents results in $2m(f + m + 1) + 2$ non-zero contributions to mixed model equations as a result of the observation, and $3(2m + 1)$ non-zero elements relating random effects for the individual to those in its parents. As a result, the total number of non-zero elements in the half-stored mixed model equations is expected to be $(2m(f + m + 4) + 5)n$. This approximation is an overestimate because elements pertaining to repeated sire dam combinations have been included and animals with unidentified parents have fewer elements. It is likely that the inverse of the combined numerator relationship matrix will contain relatively few non-zero elements and the total number of elements is $\frac{1}{2}(n + 1 + f)(n + f)$. When we ignore the f fixed effects, the number of non-zero elements in the complete model (21) will be smaller than in the model predicting the combined effect when $2m(m + 4) + 5 < \frac{1}{2}(n + 1)$ which can be approximated by $2m^2 < \frac{1}{2}n$. The combined effect model therefore requires more computer memory, and calculations per iteration. However, the combined effects model will result in equations which should converge more easily, reducing computing time.

In the above derivations we have assumed that marker genotypes are available on all animals. A more likely scenario, however, is that marker genotypes will be available on a limited number of individuals only. In this case, marker information will only affect the structure of the inverse of the combined numerator relationship matrix for some animals while for others (*e.g.*, base animals and offspring without genotypes identified) the structure will be the same as without markers. This feature makes application of the combined model attractive in

populations where marker genotypes are available on a limited number of animals. Further research on the structure of the matrices might result in reductions of computational requirements.

In this study we have assumed that a marker locus is linked to a single quantitative trait locus. The marker locus, however, might also be linked to a cluster of several genes of moderate effect (GELDERMANN 1975; DENTINE and COWAN 1990). In this case, chromosome segments rather than alleles can be followed by the segregation at the marker locus. GOLDGAR (1990) showed how the method of FERNANDO and GROSSMAN (1989), which also assumes a single MQTL, can be changed to account for a large number of QTL surrounding the marker locus.

When more than one MQTL is being considered, covariances between pairs of MQTL effects are assumed zero (Equation 23). When a trait has been under selection for some time, covariances between pairs of MQTL effects are likely to be non-zero due to linkage disequilibrium (BULMER 1985) even when they are on different chromosomes. The magnitude and sign of the covariances determine the extent of error in predicting MQTL effects due to incorrectly assuming null covariances between MQTL effects. For a trait undergoing selection, covariances are mostly negative and, as a result, MQTL effects may be overpredicted (CANTET and SMITH 1991).

In Equations 20 and 23 only the sum of all genetic effects at all MQTL and QTL are considered. This, however, might not be the optimal solution, especially not in a multiple trait situation. Selection programs in livestock are generally directed at changing more than one trait. Antagonistic relationships are generally found, for example, between production traits on the one hand and reproduction and health traits on the other. These genetic relationships can be largely explained by some loci having a positive effect on one trait and a negative effect on the other. Genetic correlations are not one, which implies that loci differ in their direction and or magnitude of effect on different traits. In such a situation we might be interested to predict the sum of effects of MQTL, which have a positive effect on one trait and no effect or a positive effect on the second trait, separately from the effects of other loci. This can be achieved by fitting two random factors in (23) for that trait. By doing this, selection pressure can be applied only to MQTL which have little or no negative effects on other traits of interest. Information from genetic maps will tell us whether this is a feasible way to achieve a well-balanced genetic progress in livestock selection.

This paper was prepared during a sabbatical leave of J.A.M.v.A. at the Department of Animal Science in Armidale, and he thanks his colleagues in Wageningen for offering him the opportunity to go on sabbatical and the Netherlands organization for Scientific Research (NWO) for their financial support. B.T. thanks the Australian Meat Research Corporation for their support.

LITERATURE CITED

BULMER, M. G., 1985 *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford.

CANTET, R. J. C., and C. SMITH, 1991 Reduced animal model for marker assisted selection using best linear unbiased prediction. Genet. Sel. Evol. **23:** 221–233.

DENTINE, M. R., and C. M. COWAN, 1990 An analytical model for the estimation of chromosome substitution effects in the offspring of individuals heterozygous at a segregating marker locus. Theor. Appl. Genet. **79:** 775–780.

FERNANDO, R. L., and M. GROSSMAN, 1989 Marker assisted selection using best linear unbiased prediction. Gen. Sel. Evol. **21:** 467–477.

FERNANDO, R. L., C. STRICKER and R. C. ELSTON, 1993 Efficient scheme to compute the posterior genotypic distribution for every member of a pedigree without loops. Theor. Appl. Genet. (in press).

GELDERMANN, H., 1975 Investigations on inheritance of quantitative characters in animals by gene markers. 1. Methods. Theor. Appl. Genet. **46:** 319–330.

GIBSON, J. P., B. W. KENNEDY, L. R. SCHAEFFER and O. I. SOUTHWOOD, 1988 Gametic models for estimation of autosomally inherited effects that are expressed only when received from either male or female parent (abstract). J. Dairy Sci. **71**(Suppl. 1): 143.

GODDARD, M. E., 1992 A mixed model for analyses of data on multiple genetic markers. Theor. Appl. Genet. **83:** 878–886.

GOLDGAR, D. E., 1990 Multipoint analysis of human quantitative variation. Am. J. Hum. Genet. **47:** 957–967.

GRASER, H.-U., S. P. SMITH and B. TIER, 1987 A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. J. Anim. Sci. **64:** 1362–1370.

HENDERSON, C. R., 1976 A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. Biometrics **32:** 69–83.

JANSS, L. L. G., J. A. M. VAN ARENDONK and J. H. J. VAN DER WERF, 1993 Identification of a major gene in several generations of data with application to F_2 and F_3 crosses. Genet. Sel. Evol. (in press).

KASHI, Y., E. HALLERMAN and M. SOLLER, 1990 Marker-assisted selection of candidate bulls for progeny testing programmes. Anim. Prod. **51:** 63–74.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

MEUWISSEN, T. H. E., and J. A. M. VAN ARENDONK, 1992 Potential improvements in rate of genetic gain from marker assisted selection in dairy cattle breeding schemes. J. Dairy Sci. **75:** 1651–1659.

QUAAS, R. L., 1976 Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics **32:** 949–953.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHAF, R. HIGUCHI *et al.*, 1988 Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. Science **239:** 487.

SCHAEFFER, L. R., B. W. KENNEDY and J. P. GIBSON, 1989 The inverse of the gametic relationship matrix. J. Dairy Sci. **72:** 1266–1272.

SMITH, S. P., and F. L. ALLAIRE, 1985 Efficient selection rules to increase nonlinear merit: application in mate selection. Genet. Sel. Evol. **17:** 387–406.

SMITH, S. P., and A. MÄKI-TANILLA, 1990 Genotypic covariance matrices and their inverses for models allowing for dominance and inbreeding. Genet. Sel. Evol. **22:** 65–91.

SOLLER, M., 1978 The use of loci associated with quantitative effects in dairy cattle improvement. Anim. Prod. **27:** 133–179.

SOLLER, M., and J. S. BECKMANN, 1983 Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. **67:** 25–33.

TIER, B., and J. SÖLKNER, 1993a Analysing gametic variation with an animal model. Theor. Appl. Genet. **85:** 868–872.

VAN ARENDONK, J. A. M., C. SMITH and B. W. KENNEDY, 1989 Method to estimate genotype probabilities at individual loci in farm livestock. Theor. Appl. Genet. **78:** 735–740.

VAN ARENDONK, J. A. M., B. TIER and B. P. KINGHORN, 1993 Simultaneous estimation of effects of unlinked markers and polygenes on a trait showing quantitative genetic variation. Proc. 17th Int. Congr. Genetics, Birmingham, UK (192).

WANG, T., S. VAN DER BEEK, R. L. FERNANDO and M. GROSSMAN, 1991 Covariance between effects of marked QTL alleles. J. Anim. Sci. **69:** 202 (supplement).

WRIGHT, S., 1922 Coefficients of inbreeding and relationship. Am. Nat. **56:** 330–338.

Communicating editor: B. S. WEIR

APPENDIX A

Derivation of vector with contributions of gametic effects (**s**) when gametic relationships are known.

Two offspring: Consider a dam without information on M, which has two offspring (i,j) with known genotype for M. Using Equation 24 the probability that the maternally derived gametic effects in i and j are identical by descent, is determined to be g_{ij}. The row and column in the gametic relationship matrix for v_i^m can be built with a vector $\mathbf{s}_i$ with two non-zero elements of $\frac{1}{2}$ corresponding to the paternal and maternal gametic effect in the dam. To determine the gametic relationships of the maternal gametic effect in the second offspring v_j^m the vector $\mathbf{s}_j$ can only have non-zero elements at the position corresponding to v_d^m, v_d^p and v_i^m, which are set equal to $s_{m,d}$, $s_{p,d}$ and $s_{m,i}$. From Equation 8 it can be seen that the following relation holds

$$g_{ij} = \tfrac{1}{2}s_{m,d} + \tfrac{1}{2}s_{p,d} + s_{m,i}. \tag{A1}$$

When there is no information on M of the dam or on its parents, the following relation holds

$$s_{m,d} = s_{p,d}. \tag{A2}$$

It is further known that

$$s_{m,d} + s_{p,d} + s_{m,i} = 1. \tag{A3}$$

From (A1), (A2) and (A3) it follows that

$$s_{m,i} = 2g_{ij} - 1$$
$$s_{m,d} = (1 - g_{ij}).$$

For the example shown in Table 9, $s_{m,i} = 0.33$ and $s_{m,d} = s_{p,d} = 0.67$.

In Equation A2 it is assumed that genotype probabilities for marker genotypes are equal for grandsire and granddam. Probabilities might be unequal when, for example, the marker genotype of the grandsire is known. In that case the relation between $s_{m,d}$ and $s_{p,d}$ will be different.

More than two offspring: In a situation where an ungenotyped animal has more than two offspring the procedure to calculate **s** starts with determining the gametic relationship matrix $\mathbf{G}_{v|r,k}$ where k is the number of gametic effect after including the offspring. The relationships in $\mathbf{G}_{v|r,k}$ which relate to the gametic effects in the offspring and those in the other animals can be calculated ignoring the marker information on the offspring. The gametic relationships between the offspring can be calculated using (24).

After constructing $\mathbf{G}_{v|r,k}$, a procedure equivalent to (11) and (12) can be used to determine the vectors **s** for the gametic effects in the offspring. From Equation 8 it can be seen that the row vector containing the gametic relationships of gametic effect k with gametic effects 1 to $k-1$ ($\mathbf{G}^i_{v|r,k}$) is equal to

$$\mathbf{G}^i_{v|r,k} = \mathbf{G}_{v|r,k-1}\mathbf{s}_k. \tag{A4}$$

Consequently, the vector $\mathbf{s}_k$ with contributions of gametic effects 1 to $k-1$ can be obtained from

$$\mathbf{s}_k = \mathbf{G}^{-1}_{v|r,k-1}\mathbf{G}^i_{v|r,k}. \tag{A5}$$

Given $\mathbf{s}_k$ we can use Equation 9 to get $\mathbf{G}^{-1}_{v|r,k}$.

The computational requirements to solve $\mathbf{s}_k$ using (A5) can be reduced considerably because only the gametic relationships pertaining to parents and all offspring have to be considered rather than the entire $\mathbf{G}^i_{v|r,k}$.