Genetic Evaluation by Best Linear Unbiased Prediction Using Marker and Trait Information in a Multibreed Population

T. Wang, R. L. Fernando and M. Grossman

Department of Animal Sciences, University of Illinois, Urbana, Illinois 61801

Manuscript received June 5, 1997
Accepted for publication September 23, 1997

ABSTRACT

Genetic evaluation by best linear unbiased prediction (BLUP) requires modeling genetic means, variances, and covariances. This paper presents theory to model means, variances, and covariances in a multibreed population, given marker and breed information, in the presence of gametic disequilibrium between the marker locus (ML) and linked quantitative trait locus (MQTL). Theory and algorithms are presented to construct the matrix of conditional covariances between relatives (Gv) for the MQTL effects in a multibreed population and to obtain the inverse of Gv efficiently. Theory presented here accounts for heterogeneity of variances among pure breeds and for segregation variances between pure breeds. A numerical example was used to illustrate how the theory and algorithms can be used for genetic evaluation by BLUP using marker and trait information in a multibreed population.

DEFINITION AND NOTATION

A multibreed population comprises individuals from several breed groups, including purebred and crossbred groups (Lo et al. 1993). Founders are assumed to be unrelated, noninbred, and of known breed; parents of nonfounders are assumed to be known. Breed information, denoted B, consists of breeds for founders and parentage for nonfounders. Breed information (B) for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects for the pedigree in Table 1, for example, consists of breed groups (L, P, o) for the MQTL effects. Assume gametic equilibrium between the marker locus (ML) and linked quantitative trait locus (MQTL), and assume the genotype at the ML is known for each individual in the pedigree. Following Wang et al. (1995), for individual i, let \( M_i \) and \( M' \) denote the two alleles at the ML, and let \( Q_i \) and \( Q' \) denote the two alleles at the MQTL linked to \( M_i \) and \( M' \):

\[
\begin{array}{c}
M_i \\
M' \\
Q_i \\
Q'\
\end{array}
\]

Further, let \( v_i^1 \) and \( v_i^2 \) be the additive effects of \( Q_i \) and \( Q' \) in individual i, and let \( u_i \) be the sum of additive effects of the remaining QTL alleles (RQTL). Assume
The RQTL are unlinked and in genetic equilibrium with the MQTL and with the ML; thus, \( u_i \) is uncorrelated with \( v_i^1 \) and with \( v_i^2 \). Let \( v \) be a vector of \( v_i^1 \) and \( v_i^2 \) for \( i = 1, 2, \ldots, n \), and \( u \) be a vector of \( u_i \) for \( i = 1, 2, \ldots, n \). Suppose \( J \) alleles segregate at the MQTL in a multibreed population with allele frequencies \( p_{ij} \), \( i = 1, \ldots, J \) in purebred group \( B_i \). Then, \( v_i^1 \) and \( v_i^2 \) are random variables with sample space \( \{ v_{ij} \) : \( j = 1, \ldots, J \} \), where \( v_j \) is the effect of MQTL allele \( j \) in some reference breed group (e.g., one of the pure breeds). Differences in allele frequencies among breed groups give rise to heterogeneous means, variances, and covariances.

To model means, variances, and covariances of additive genetic values \( (a_i) \), let \( a_i = v_i^1 + v_i^2 + u_i \). In matrix notation, the vector of additive genetic values \( (a) \) is

\[
a = Lv + u
\]

(1)

where

\[
L = \begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix}
\]

is a matrix of order \( n \times 2n \). Now, the conditional mean of \( a \), given breed \( (B) \) and marker \( (M) \) information, can be written as \( E(a | BM) = LE(v | BM) + E(u | BM) \). As already mentioned, RQTL are unlinked with the ML, so that \( E(u | BM) = E(u | B) \) and

\[
E(a | BM) = LE(v | BM) + E(u | B)
\]

(2)

The second term on the right-hand side of (2) depends only on breed information, so it can be modeled using genetic group theory. The first term, however, depends on breed and marker information, and the modeling of this term is presented here.

The \( u_i \) is uncorrelated with both \( v_i^1 \) and \( v_i^2 \), so that the conditional variance and covariance matrix of \( a \), given breed and marker information, can be written as

\[
\text{Var}(a | BM) = LGvL' + G_u
\]

(3)

where \( G_v \) denotes \( \text{Var}(v | BM) \) and \( G_u \) denotes \( \text{Var}(u | BM) \).

The second term on the right-hand side of (3) depends only on breed information, so it can be modeled using additive covariance theory for a multibreed population (Lo et al. 1993).

**Means of MQTL Effects**

Let \( i \) be an individual with sire \( s \) and dam \( d \) (Figure 1). The conditional expectation of the MQTL effect \( v_i^k \) for \( k = 1, 2 \), given breed \( (B) \) and marker \( (M) \) information, can be written as \( E(v_i^k | BM) = E(EM_i^k | Q_i^k \leftarrow B_iBM) \), where \( Q_i^k \leftarrow B \), means that \( Q_i^k \) can be traced back to pure breed \( B_i \), or that \( Q_i^k \) originates from pure breed \( B_i \). Then,

\[
E(v_i^k | BM) = \sum_{l=1}^{L} E(v_i^k | Q_i^k \leftarrow B_iBM) \Pr(Q_i^k \leftarrow B_iBM) = \sum_{l=1}^{L} \mu_{B_i}^k f_{B_i}^k
\]

(4)

where \( \mu_{B_i}^k = E(v_i^k | Q_i^k \leftarrow B_iBM) = \sum_{l=1}^{L} p_{i_j^l}^j v_{ij} \) is the mean of additive MQTL effects in pure breed \( B_i \), and \( f_{B_i}^k = \Pr(Q_i^k \leftarrow B_iBM) \) is the conditional probability that MQTL allele \( Q_i^k \) originates from pure breed \( B_i \), given breed and marker information. This probability will be also referred to as the conditional breed composition for MQTL allele \( Q_i^k \), given breed and marker information.

Note that the expected value of \( v_i^k \) for any individual \( i \) in a multibreed population, given breed and marker information, can be written as a linear combination of the means \( (\mu_{B_i}^k) \) of additive MQTL effects in pure breeds.

Computation of \( f_{B_i}^k \) in (4) is described below. Recall that \( s \) and \( d \) are the parents of individual \( i \), so that MQTL allele \( Q_i^k \) can originate from a pure breed \( B_i \) allele in one of four ways:

(1) \( Q_i^k \) descends from \( Q_s^1 \), and \( Q_i^1 \) originated from pure breed \( B_s \); denoted \( Q_i^k \leftarrow Q_s^1, Q_s^1 \leftarrow B_s \)

(2) \( Q_i^k \) descends from \( Q_d^2 \), and \( Q_i^1 \) originated from pure breed \( B_d \); denoted \( Q_i^k \leftarrow Q_d^2, Q_d^2 \leftarrow B_d \)

\[
\begin{array}{c c c c c c}
M_s^1 & Q_s^1 & & & & \\
M_s^2 & Q_s^2 & & & & \\
\downarrow & & & & & \\
M_d^1 & Q_d^1 & & & & \\
M_d^2 & Q_d^2 & & & & \\
\end{array}
\]

Figure 1.—Chromosomal fragments containing the ML and the MQTL for individuals \( s, d, i, \) and \( j \) where \( s \) and \( d \) are parents of \( i \) and where \( j \) is not a direct descendant of \( i \).
(3) \( Q_k^i \) descended from \( Q_3^i \), and \( Q_2^i \) originated from breed \( B_i \), denoted \( (Q_k^i \leftarrow Q_3^i, Q_2^i \leftarrow B_i) \)

(4) \( Q_k^i \) descended from \( Q_3^i \), and \( Q_2^i \) originated from breed \( B_i \), denoted \( (Q_k^i \leftarrow Q_3^i, Q_2^i \leftarrow B_i) \)

where, for example, \( Q_3^i \) and \( Q_2^i \) are the MQTL alleles in parent \( s \). The \( f_{B_i}^{i_1,i_2} \), therefore, can be written as

\[
f_{B_i}^{i_1,i_2} = \text{Pr}(Q_{i_1}^{i_1} \leftarrow B_i | BM) = \text{Pr}(Q_{i_1}^{i_1} \leftarrow Q_3^i, Q_2^i \leftarrow B_i | BM) + \text{Pr}(Q_{i_1}^{i_1} \leftarrow Q_3^i, Q_2^i \leftarrow B_i | BM)
\]

Note that the sampling of \( Q_k^i \) from parents \( s \) and \( d \) does not depend on the breed origin of the MQTL alleles in \( s \) and \( d \). Thus, (5) can be computed recursively as

\[
f_{B_i}^{i_1,i_2} = \text{Pr}(Q_{i_1}^{i_1} \leftarrow B_i | BM)
\]

where \( \text{Pr}(Q_{i_1}^{i_1} \leftarrow Q_3^i | BM) \) does not depend on breed information, so \( \text{Pr}(Q_{i_1}^{i_1} \leftarrow Q_3^i | BM) = \text{Pr}(Q_{i_1}^{i_1} \leftarrow Q_3^i | M) \). Wang et al. (1995) referred to this as the probability of descent for MQTL alleles (PDQ) and presented a recursive formula for its computation.

**VARIANCE OF MQTL EFFECTS**

Given random variables \( X \) and \( Y \), the variance of \( X \) can be written as

\[
\text{Var}(X) = \text{E}[\text{Var}(X \mid Y)] + \text{Var}[\text{E}(X \mid Y)]
\]

(Kempthorne and Folsom 1971). Using the above formula, the conditional variance of MQTL effect \( v_k^i \) for \( k = 1 \) or \( 2 \), given breed and marker information, can be written as

\[
\text{Var}(v_k^i | BM) = \text{E}[\text{Var}(v_k^i | Q_k^i \leftarrow B_i, BM)] + \text{Var}[\text{E}(v_k^i | Q_k^i \leftarrow B_i, BM)]
\]

The first term of (8) can be expressed in terms of conditional breed compositions as

\[
E[\text{Var}(v_k^i | Q_k^i \leftarrow B_i, BM)] = \sum_{i=1}^{L} \text{Var}(v_k^i | Q_k^i \leftarrow B_i, BM) \text{Pr}(Q_k^i \leftarrow B_i | BM) = \sum_{i=1}^{L} \sigma_{v_k^i}^{2} f_{B_i}^{i,i_k}
\]

where \( \sigma_{v_k^i}^{2} = \text{Var}(v_i^k | Q_k^i \leftarrow B_i, BM) = \sum_{l=1}^{L} (v_l - \mu_{B_i})^2 p_{B_i}^{l} \) is the variance of the MQTL effects in pure breed \( B_i \). Note that the quantity on the right-hand side of (9) is the mean of variances \( \sigma_{v_k^i}^{2} \) of the MQTL effects in the pure breeds, weighted by conditional breed compositions. Purebred variances \( \sigma_{v_k^i}^{2} \) for \( l = 1, \ldots, L \), can be estimated by maximum likelihood (Weiler and Fernando 1991).

The second term of (8) also can be expressed in terms of conditional breed composition as

\[
\text{Var}[\text{E}(v_k^i | Q_k^i \leftarrow B_i, BM)] = \sum_{l=1}^{L} \text{E} \left( \sum_{i=1}^{L} \text{Var}(v_k^i | Q_k^i \leftarrow B_i, BM) \right)
\]

\[
= \sum_{l=1}^{L} \sum_{i=1}^{L} (\mu_{B_l} - \mu_{B_i})^2 f_{B_l}^{i,i_k} f_{B_i}^{i,k}
\]

**COVARIANCE BETWEEN RELATIVES FOR MQTL EFFECTS**

Following Wang et al. (1995), conditional covariances between relatives for additive effects of MQTL alleles, given marker and breed information, will be derived separately for alleles between the two individuals and for alleles within an individual.

**Between individuals**: Let \( i \) be an individual with sire \( s \) and dam \( d \), and let \( j \) be another individual who is not a direct descendant of \( i \) (Figure 1). Given random variables \( X, Y, \) and \( Z \), the covariance between \( X \) and \( Y \) can be written as

\[
\text{Cov}(X, Y) = \text{E}[\text{Cov}(X, Y | Z)] + \text{Cov}[\text{E}(X | Z), \text{E}(Y | Z)]
\]

(Kempthorne and Folsom 1971). Using the above formula, the conditional covariance between relatives for additive MQTL effects \( v_k^i \) and \( v_k^j \), given breed and marker information, can be written as

\[
\text{Cov}(v_k^i, v_k^j | BM) = \text{E}[\text{Cov}(v_k^i, v_k^j | BM, W)] + \text{ Cov}[\text{E}(v_k^i | BM, W), \text{E}(v_k^j | BM, W)]
\]

(13)
where $W$ is a random variable with sample space \{$Q_1^i \equiv Q_0^i, Q_1^i \equiv Q_2^i, Q_2^i \equiv Q_3^i, Q_3^i \equiv Q_1^i$\}. Note that $v_i^j$ is independent of $W$ because $j$ is not a direct descendant of $i$ (Chang et al. 1991). Thus, $E(v_i^j | BM,W)$ does not depend on $W$: $E(v_i^j | BM,W) = E(v_i^j | BM)$. Therefore, the second term of (13) is null.

The first term of (13) can be expressed in terms of PDQs (Wang et al. 1995):

$$E[\text{Cov}(v_i^j, v_j^k | BM, W)]$$

$$= \text{Cov}(v_i^j, v_j^k | Q_i^k \equiv Q_1^j, BM) \cdot \text{Pr}(Q_i^k \equiv Q_1^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | Q_i^k \equiv Q_2^j, BM) \cdot \text{Pr}(Q_i^k \equiv Q_2^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | Q_i^k \equiv Q_3^j, BM) \cdot \text{Pr}(Q_i^k \equiv Q_3^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | Q_i^k \equiv Q_4^j, BM) \cdot \text{Pr}(Q_i^k \equiv Q_4^j | BM)$$

The $\text{Cov}(v_i^j, v_j^k | Q_i^k \equiv Q_1^j, BM)$ can be computed recursively as

$$\text{Cov}(v_i^j, v_j^k | BM) = \text{Cov}(v_i^j, v_j^k | BM) \cdot \text{Pr}(Q_i^k \equiv Q_1^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | BM) \cdot \text{Pr}(Q_i^k \equiv Q_2^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | BM) \cdot \text{Pr}(Q_i^k \equiv Q_3^j | BM)$$

$$+ \text{Cov}(v_i^j, v_j^k | BM) \cdot \text{Pr}(Q_i^k \equiv Q_4^j | BM)$$

(14)

This recursive formula is identical to that used for purged populations, given marker information (Hoeschele et al. 1993; Wang et al. 1995).

**Within an individual:** For individual $i$, we present the computation of the covariance between additive effects $v_i^1$ and $v_i^2$ of MQT alleles $Q_1^i$ and $Q_2^i$.

Following Mendelian inheritance, the pair of homologous alleles at the MQT, $Q_1^i$ and $Q_2^i$, descended from one of the following pairs of parents: $(Q_1, Q_2), (Q_2, Q_1), (Q_1, Q_2)$, or $(Q_2, Q_1)$. Let $E_{kkd}$ denote the event that the pair of alleles in $i$ descended from parents $Q_1^i \equiv Q_0^i, Q_2^i \equiv Q_3^i$ or $Q_2^i \equiv Q_0^i, Q_3^i \equiv Q_1^i$, for $k = 1$ or 2. Using the formula (12), the covariance between relatives for additive MQL effects $v_i^1$ and $v_i^2$ can be written as

$$\text{Cov}(v_i^1, v_i^2 | BM) = E[\text{Cov}(v_i^1, v_i^2 | BM, W)]$$

$$+ \text{Cov}(E(v_i^1 | BM, W), E(v_i^2 | BM, W))$$

(15)

where $W$ is a random variable with sample space $\{E_{kkd} \}$ for $k = 1$ or 2. The first term in (15) is

$$E[\text{Cov}(v_i^1, v_i^2 | BM, W)] = \sum_{k=1}^{2} \sum_{k_d=1}^{2} \text{Cov}(v_{s, k}^1, v_{d, k}^2 | BM) \cdot \text{Pr}(E_{kkd} | BM)$$

(16)

Note that event $E_{kkd}$ does not depend on breed information, so that $\text{Pr}(E_{kkd} | BM) = \text{Pr}(E_{kkd} | M)$. Following Wang et al. (1995), $\text{Pr}(E_{kkd} | M)$ can be computed as

$$\text{Pr}(E_{kkd} | M) = \text{Pr}(Q_i^1 \equiv Q_0^i, Q_i^2 \equiv Q_3^i | M) + \text{Pr}(Q_i^1 \equiv Q_3^i, Q_i^2 \equiv Q_0^i | M)$$

$$= \frac{\text{Pr}(Q_i^1 \equiv Q_0^i | M) \cdot \text{Pr}(Q_i^2 \equiv Q_3^i | M) \cdot \text{Pr}(Q_i^1 \equiv Q_3^i | M) \cdot \text{Pr}(Q_i^2 \equiv Q_0^i | M)}{\text{Pr}(Q_i^1 \equiv Q_0^i | M) + \text{Pr}(Q_i^1 \equiv Q_3^i | M) + \text{Pr}(Q_i^2 \equiv Q_0^i | M) + \text{Pr}(Q_i^2 \equiv Q_3^i | M)}$$

(17)

The second term in (15) is

$$\text{Cov}[E(v_i^1 | BM, W), E(v_i^2 | BM, W)] =$$

$$E[\text{Cov}(v_i^1 | BM, W)E(v_i^2 | BM, W)] - E[(v_i^1 | BM)(v_i^2 | BM)]$$

(18)

where $E[(v_i^1 | BM, W) E(v_i^2 | BM, W)] = \sum_{k=1}^{2} \sum_{k_d=1}^{2} E(v_{s, k}^1 | BM) E(v_{d, k}^2 | BM) \cdot \text{Pr}(E_{kkd} | M)$. Therefore, substituting (16) and (18) into (15), we have

$$\text{Cov}(v_i^1, v_i^2 | BM) = \sum_{k=1}^{2} \sum_{k_d=1}^{2} \text{Cov}(v_{s, k}^1, v_{d, k}^2 | BM)$$

$$+ E(v_{s, k}^1 | BM) E(v_{d, k}^2 | BM) \cdot \text{Pr}(E_{kkd} | M)$$

$$- E(v_i^1 | BM) E(v_i^2 | BM)$$

(19)

In (19), $\text{Cov}(v_{s, k}^1, v_{d, k}^2 | BM)$ is computed recursively using (14), and means of additive MQTL effects $E(v_i^1 | BM), E(v_i^2 | BM)$, $E(v_i^1 | BM)$, and $E(v_i^2 | BM)$ are computed using (4).

**Tabular Method to Compute Covariance Matrix $G_i$:**

Following Wang et al. (1995), we show how (11), (14), and (19) can be used to construct recursively the covariance matrix $G_i$ between relatives for the MQTL effects by a tabular method.

For each individual in the pedigree, matrix $G_i$ included a row and a column for each of the two MQTL alleles. Thus, the order of $G_i$ is $2n$, where $n$ is the number of individuals. Each diagonal element of $G_i$ is given by (11). Individuals are ordered such that parents pre-
cede their progeny, and individuals 1 through b are considered to be unrelated and noninbred. Thus, the upper left submatrix of \( G_b \) is a diagonal matrix of order 2b, which will be expanded sequentially by the two rows and the two columns corresponding to individual i, for \( i = b + 1 \), \ldots, n.

Let \( G_{i-1} \) be the upper left submatrix of \( G_b \) expanded up through individual \( i - 1 \). For individual \( i \), with parents \( s \) and \( d \), \( G_{i-1} \) is expanded to \( G_b \) as

\[
G_i = \begin{bmatrix} G_{i-1} & G_{i-1}q_i \\ q_i' & G_{i-1}C_i \end{bmatrix}
\]

where

\[
C_i = \begin{bmatrix} \text{Var}(v_i'BM) & \text{Cov}(v_i'v_i'BM) \\ \text{Cov}(v_i'v_i'BM) & \text{Var}(v_i'BM) \end{bmatrix}
\]

and

\[
q_i' = \begin{bmatrix} 0 & \cdots & 0 & T_i(1,1) & T_i(1,2) & \cdots & T_i(1,3) & T_i(1,4) & \cdots & 0 & 0 \\ 0 & \cdots & 0 & T_i(2,1) & T_i(2,2) & \cdots & T_i(2,3) & T_i(2,4) & \cdots & 0 & 0 \end{bmatrix}
\]

Matrix \( q_i' \) is of order 2 \( \times \) (\( i - 1 \)), with at most eight nonzero elements that are from:

\[
T_i = \begin{bmatrix} \text{Pr}(Q_i = 0_1M) & \text{Pr}(Q_i = 0_2M) & \text{Pr}(Q_i = 0_3M) & \text{Pr}(Q_i = 0_4M) \\ \text{Pr}(Q_i = 0_5M) & \text{Pr}(Q_i = 0_6M) & \text{Pr}(Q_i = 0_7M) & \text{Pr}(Q_i = 0_8M) \end{bmatrix}
\]

These nonzero elements in \( q_i' \) are located in columns \( 2(s - 1) + 1, 2(s - 1) + 2, 2(d - 1) + 1, \) and \( 2(d - 1) + 2 \). Diagonal elements in (21) are computed using (11) and off-diagonal elements are computed using (19).

**Efficient Algorithm to Invert Covariance Matrix \( G_v \)**

The \( q_i' \) in (22) is sparse, and positions of the nonzero elements can be determined easily; therefore, \( G_b \) can be inverted efficiently (van Arendonk et al. 1994; Wang et al. 1995).

Following Wang et al. (1995), three matrices will be used to describe the rules to invert \( G_b \): \( D_i \), \( W_i \), and \( \Pi_i \). The first matrix is

\[
D_i = C_{1-1}T_iC_{sd}T_i'
\]

where \( C_{sd} \) is the \( 4 \times 4 \) matrix of conditional covariances between \( s \) and \( d \), the parents of \( i \), for additive MQTL effects, and \( T_i \) is the matrix of PDQs defined in (23). The second matrix is

\[
W_i = \begin{bmatrix} -T_i' & I_2 \end{bmatrix}D_i^{-1}\begin{bmatrix} -T_i & I_2 \end{bmatrix}
\]

where \( I_2 \) is an identity matrix of order 2. The third matrix is

\[
\Pi_i = \begin{bmatrix} (\delta_1, \delta_2) & (\delta_2, \delta_1) & (\delta_3, \delta_4) & (\delta_4, \delta_3) & (\delta_5, \delta_6) & (\delta_6, \delta_5) \\ (\delta_2, \delta_1) & (\delta_3, \delta_4) & (\delta_4, \delta_3) & (\delta_5, \delta_6) & (\delta_6, \delta_5) \\ (\delta_3, \delta_4) & (\delta_4, \delta_3) & (\delta_1, \delta_2) & (\delta_2, \delta_1) & (\delta_5, \delta_6) & (\delta_6, \delta_5) \\ (\delta_4, \delta_3) & (\delta_5, \delta_6) & (\delta_6, \delta_5) & (\delta_1, \delta_2) & (\delta_2, \delta_1) \\ (\delta_5, \delta_6) & (\delta_6, \delta_5) & (\delta_1, \delta_2) & (\delta_2, \delta_1) & (\delta_3, \delta_4) & (\delta_4, \delta_3) \\ (\delta_6, \delta_5) & (\delta_5, \delta_6) & (\delta_4, \delta_3) & (\delta_3, \delta_4) & (\delta_1, \delta_2) & (\delta_2, \delta_1) \\ 
\end{bmatrix}
\]

where \( \delta_i = 2(a - 1) + b \) for \( a = s, d, \) or \( i, \) and \( b = 1 \) or 2. If both parents of individual \( i \) are known, then all elements in \( \Pi_i \) are defined. If at least one parent is unknown, then elements in \( \Pi_i \) associated with the unknown parent(s) are not defined.

Rules to invert \( G_b \) are as follows: (1) Set \( G_{i-1}^{-1} \) equal to the null matrix of order 2n. (2) For individual \( i, i = 1, \ldots, n, \) (a) if both parents are unknown, then (i) add \( \text{Var}(v_i'BM) \) to \( G_{i-1}^{-1} \) (\( \delta_i \delta_j \)) and (ii) add \( \text{Var}(v_i'BM) \) to \( (\delta_i \delta_j) \); (b) if at least one parent is known, then (i) compute \( T_i \) according to (23), (ii) compute \( D_i \) according to (24), (iii) compute \( W_i \) according to (25), and (iv) for each “defined” element \( (l, k) \) in \( \Pi_i \), add element \( W_i(l, k) \) to \( G_{i-1}^{-1} \) at the position given by \( \Pi_i(l, k) \).

**Numerical Example**

To illustrate how the theory and algorithms in this paper can be used for genetic evaluation by BLUP using marker and trait information in a multibred population, we use the pedigree in Table 1. The pedigree consists of six individuals from a multibred population with two purebreds groups, \( B_1 \) and \( B_2 \), and two crossbred groups, \( B_1 \times B_2 \) and \( B_1 \times B_2 \). Individuals are ordered such that parents precede their progeny.

The mixed linear model

\[
y = X_u\beta_u + X_v\beta_v + Z_uu^* + Z_vv^* + e
\]

will be used for genetic evaluation by BLUP, where \( \beta_u \) is a vector of purebred means for additive MQTL effects, \( \beta_v \) is a vector of the fixed effects, \( v^* \) and \( u^* \) are vectors of additive MQTL and RQTL effects deviated from their respective means, \( e \) is a vector of residuals, and \( X_u, X_v, Z_u, \) and \( Z_v \) are incidence matrices. For simplicity, assume that \( \beta_v \) contains only means of RQTL effects for pure breeds \( B_1 \) and \( B_2 \), i.e., \( \beta_v = [\nu_{b1}, \nu_{b2}] \). From the definition of \( v^* \) and \( u^* \), it follows that each of \( v^* \) and \( u^* \) has null mean and that \( \text{Var}(v^*) = \text{Var}(u^*) = (G_v) \text{Var}(u^*) = \text{Var}(u) = (G_u) \). Assume also that the residual \( e \) has null mean and variance \( \sigma_e^2 \).

For this numerical example, the following hypothetical values will be used: \( \nu_{b1} = -0.1, \nu_{b2} = -0.3, \sigma_{b1}^2 = 0.29, \sigma_{b2}^2 = 0.81, \sigma_{b3}^2 = 1.2, \sigma_{b4}^2 = 1.6, \sigma_e^2 = 0.9, \sigma_e^2 \)
= 1.0. The estimation of these parameters is discussed in the discussion.

The best linear unbiased estimate (BLUE) of \( \beta_u \) and BLUP of \( \mathbf{v}^* \) and of \( \mathbf{u}^* \) can be obtained by solving the mixed model equations (Henderson 1973) given below

\[
\begin{bmatrix}
\mathbf{X}_u' \mathbf{X}_u & \mathbf{X}_u' \mathbf{Z}_u & \\
\mathbf{Z}_u' \mathbf{X}_u & \mathbf{Z}_u' \mathbf{Z}_u + \mathbf{G}_v^{-1}/\sigma_e^2 & \mathbf{Z}_u' \mathbf{Z}_u \\
\end{bmatrix}
\begin{bmatrix}
\beta_u \\
\mathbf{v}^* \\
\mathbf{u}^* \\
\end{bmatrix}
= \begin{bmatrix}
\mathbf{X}_u' \mathbf{y} \\
\mathbf{Z}_u' \mathbf{y} \\
\end{bmatrix}
\]

(28)

Construction of the incidence matrices and covariance matrices in (28) are presented in the following sections. To construct \( \mathbf{X}_u \) and \( \mathbf{G}_v \) requires conditional breed compositions \( f_{1,b}^{i,k} \), and their computations are discussed next.

**Computing conditional breed compositions \( f_{1,b}^{i,k} \):** Individuals 1 and 2 are founders from pure breed \( B_1 \), and 3 is a founder from pure breed \( B_2 \). Thus, their breed compositions are

\[
\begin{bmatrix}
0.25 & 0.25 \\
0.25 & 0.25 \\
\end{bmatrix}
= \begin{bmatrix}
1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
0.5 & 0.5 \\
0.5 & 0.5 \\
\end{bmatrix}
= \begin{bmatrix}
0.75 & 0.25 \\
0.25 & 0.75 \\
\end{bmatrix}
\]

For nonfounders 4, 5, and 6, their conditional breed compositions can be computed using recursive formula (6), given PDQs. The PDQs for 4, 5, and 6 (Wang et al. 1995) are

\[
T_4 = \begin{bmatrix}
0.25 & 0.25 & 0.25 & 0.25 \\
0.25 & 0.25 & 0.25 & 0.25 \\
\end{bmatrix}
\]

\[
T_5 = \begin{bmatrix}
0.05 & 0.45 & 0.05 & 0.45 \\
0.45 & 0.05 & 0.45 & 0.05 \\
\end{bmatrix}
\]

and

\[
T_6 = \begin{bmatrix}
0.25 & 0.25 & 0.25 \\
0.25 & 0.25 & 0.25 \\
0.5 & 0.5 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\]

These PDQs are used now to compute conditional breed compositions \( f_{1,b}^{i,k} \), for \( i = 4, 5, 6 \), using (6):
Matrix $Z_r = I_6 \otimes J_2$ is the Kronecker product of an identity matrix of order 6 and a $2 \times 1$ vector of ones. Matrix $Z_r$ is an identity matrix of order 6.

Constructing $G_5$: Recall that individuals 1 and 2 are founders from pure breed $B_1$ and 3 is a founder from pure breed $B_2$ and that founders are unrelated and noninbred. Thus, the upper-left submatrix of $G_5$ is diagonal and of order 6, with diagonal elements $\alpha_{vB_1}^2$, $\alpha_{vB_1}^2$, $\alpha_{vB_2}^2$, $\alpha_{vB_2}^2$, and $\alpha_{vB_1}^2$. For nonfounders 4, 5, and 6, this upper-left submatrix can be expanded by the tabular method using (20).

For individual 4, the matrix $q'_4$ for use in (20) is

$$q'_4 = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0 & 0 \end{bmatrix}$$

The first six elements in rows 7 and 8 of $G_5$ are computed as $q'_4 G_3$ (Table 2); columns 7 and 8 are obtained by symmetry. Each of the two diagonal elements of $C_4$ is computed using (11). Individual 4 is from pure breed $B_1$; therefore, each diagonal of $C_4$ is $\sigma_{vB_1}^2$. Parents of 4 are unrelated, so off-diagonal elements of $C_4$ are null. Thus,

$$C_4 = \begin{bmatrix} 0.29 & 0 \\ 0 & 0.29 \end{bmatrix}$$

For individual 5, the matrix $q'_5$ for use in (20) is

$$q'_5 = \begin{bmatrix} 0 & 0.05 & 0.45 & 0.05 & 0.45 & 0 & 0 \\ 0 & 0.45 & 0.05 & 0.45 & 0 & 0 \end{bmatrix}$$

Each diagonal element of $C_5$ is computed using (11): $\text{Var}(v_{B_1} | BM) = (0.29)(0.5) + (0.81)(0.5) + (-0.1 - (-0.3))(0.5)(0.5) = 0.559$. The Var($v_{B_1} | BM$) happens to have the same value as $\text{Var}(v_{B_1} | BM)$. Parents of 5 are unrelated, so off-diagonal elements of $C_5$ are null. Thus,

$$C_5 = \begin{bmatrix} 0.559 & 0 \\ 0 & 0.559 \end{bmatrix}$$

For individual 6, the matrix $q'_6$ for use in (20) is

$$q'_6 = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.00 & 0.00 & 0.00 & 0.50 & 0.50 & 0.00 & 0.00 \end{bmatrix}$$

Parents of 6 are related, so now we have to compute the conditional covariance $\text{Cov}(v_{B_1} | v_{B_2} | BM)$, which requires $\text{Cov}(v_{B_1} | v_{B_2} | BM)$, $\text{Pr}(k_{4k55} | M)$, $\text{E}(v_{B_1} | BM)$, and $\text{E}(v_{B_2} | BM)$ for $k_4$, $k_5$, $k_6 = 1$ or 2. The $\text{Cov}(v_{B_1} | v_{B_2} | BM)$ for $k_4$, $k_5 = 1$ or 2 have already been computed as $\text{Cov}(v_{B_1} | v_{B_2} | BM) = \text{Cov}(v_{B_1} | v_{B_2} | BM) = \text{Cov}(v_{B_1} v_{B_2} | BM) = 0.036$. The Pr $(E_{k4k55} | M)$ can be computed using (17); for example,

$$\text{Pr}(E_{11} | M) = \frac{T_6(1,1)T_6(2,3) + T_6(1,3)T_6(2,1)}{T_6(1,1) + T_6(1,2) + T_6(1,3) + T_6(1,4)} = 0.25$$

Similarly, $\text{Pr}(E_{12} | M) = \text{Pr}(E_{21} | M) = \text{Pr}(E_{22} | M) = 0.25$. The conditional expected values of additive MQTL effects, $\text{E}(v_{B_1} | BM)$, $\text{E}(v_{B_2} | BM)$, and $\text{E}(v_{B_2} | BM)$, are computed using (4) as

$$\text{E}(v_{B_1} | BM) = \text{E}(v_{B_2} | BM) = (0.1)(1.0) + (0.3)(0.0) = -0.1$$

$$\text{E}(v_{B_2} | BM) = \text{E}(v_{B_2} | BM) = (0.1)(0.5) + (0.3)(0.5) = -0.2$$

$$\text{E}(v_{B_2} | BM) = (0.1)(0.75) + (0.3)(0.25) = -0.15$$

$$\text{E}(v_{B_2} | BM) = (0.1)(1.00) + (0.3)(0.00) = -0.1$$

---

### Table 2

The conditional covariance matrix ($G_n$) for MQTL effects

<table>
<thead>
<tr>
<th></th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
<th>$v_6$</th>
<th>$v_7$</th>
<th>$v_8$</th>
<th>$v_9$</th>
<th>$v_{10}$</th>
<th>$v_{11}$</th>
<th>$v_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_1$</td>
<td>0.290</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>$v_2$</td>
<td>0.000</td>
<td>0.290</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>$v_3$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.290</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>$v_4$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.290</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.036</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>$v_5$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.810</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.045</td>
<td>0.102</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_6$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.365</td>
<td>0.045</td>
<td>0.102</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_7$</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.012</td>
<td>0.145</td>
<td>0.145</td>
</tr>
<tr>
<td>$v_8$</td>
<td>0.036</td>
<td>0.036</td>
<td>0.073</td>
<td>0.073</td>
<td>0.102</td>
<td>0.012</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_9$</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_{10}$</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_{11}$</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$v_{12}$</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.073</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Therefore,
\[
\text{Cov}(\mathbf{v}_b^2|\mathbf{v}_b^2|) = \sum_{k_1=1}^{2} \sum_{k_2=1}^{2} \mathbf{Cov}(\mathbf{v}_{k_1}^2|\mathbf{v}_{k_2}^2|) \mathbf{B} \mathbf{M} = \mathbf{E}(\mathbf{v}_{k_1}^2|\mathbf{v}_{k_2}^2|) \mathbf{E}(\mathbf{v}_{k_1}^2|\mathbf{v}_{k_2}^2|)
\]

Diagonal elements of \( \mathbf{C}_b \) are computed using (11) as

\[
\text{Var}(\mathbf{v}_b^2|\mathbf{B}\mathbf{M}) = (0.29)(0.75) + (0.81)(0.25)
\]

\[
+ (-0.1 - (-0.3))^2(0.75)(0.25) = 0.4275
\]

\[
(\text{Var}(\mathbf{v}_b^2|\mathbf{B}\mathbf{M}) = (0.29)(1.00) + (0.81)(0.00)
\]

\[
+ (-0.1 - (-0.3))^2(1.00)(0.00) = 0.29
\]

Thus,

\[
\mathbf{C}_b = \begin{bmatrix}
0.4275 & 0.041 \\
0.041 & 0.29
\end{bmatrix}
\]

The complete \( \mathbf{G}_u \) is in Table 2.

BLUE and BLUP: The covariance matrix \( \mathbf{G}_u \) for additive RQTL effects (lo et al. 1993) is

\[
\mathbf{G}_u = \begin{bmatrix}
1.2 & 0 & 0 & 0.6 & 0.3 \\
0 & 1.2 & 0 & 0.6 & 0.6 \\
0 & 0 & 1.6 & 0.8 & 0.4 \\
0.6 & 0.6 & 0.1 & 0.3 & 0.75 \\
0 & 0.6 & 0.8 & 1.4 & 0.85 \\
0.3 & 0.6 & 0.4 & 0.75 & 1.9
\end{bmatrix}
\]

Now all elements in (28) are available. The BLUE of \( \mathbf{u}_s (\mathbf{P}_s) \) and the BLUP of \( \mathbf{v}^* (\mathbf{V}^*) \) and of \( \mathbf{u}^* (\mathbf{U}^*) \) were obtained by solving (28):

\[
\hat{\mathbf{P}}_s = [11.3049, 13.6325]
\]

\[
\hat{\mathbf{v}}^* = [0.0547, 0.0547, -0.0277, -0.0277, -0.0695, -0.0695, 0.2773, 0.2773, -0.1570, -0.1570, 0.2307, 0.3404]
\]

\[
\hat{\mathbf{u}}^* = [0.12050, -0.12050, 0.0000, 1.0023, -0.2094, 1.1275]
\]

DISCUSSION

Linear-model methods use only information contained in means, variances, and covariances. When there is gametic disequilibrium between a marker locus (ML) and a closely linked QTL (MQTL), differences between means and between variances provide information on the association between the ML and the MQTL, but differences between covariances do provide information on this association. Some methods that have been proposed for ML-MQTL studies use only information contained in means (Land e and Thompson 1990; Weigel et al. 1990; Zhang and Smith 1992; Haley et al. 1994). Other methods use information contained in covariances (Fernando and Grossman 1989; Weller et al. 1990; Hoeschele 1993; Hoeschele and VanRad en 1993; Wang et al. 1995; van Arendonk et al. 1994). The method proposed in this paper for genetic evaluation by BLUP uses all information contained in means, variances, and covariances.

The conditional covariance matrix \( \mathbf{G}_u \) is a function of purebred means and variances for additive MQTL effects, and \( \mathbf{G}_u \) is a function of purebred variances and segregation variances for RQTL effects (Lo et al. 1993). Thus, genetic evaluation by BLUP requires purebred means \( \mathbf{u}_s \) and variances \( \sigma^2_{\mathbf{u},s} \) for additive MQTL effects and purebred variances \( \sigma^2_{\mathbf{u},s} \) and segregation variance for RQTL effects \( \sigma^2_{\mathbf{r},s} \). These parameters can be established by maximum likelihood, in principle, based on the theory and algorithms presented here.

For genetic evaluation by BLUP using only trait information in multibreed populations, El zo (1990) and Lo et al. (1993) presented a theory to account for heterogeneity of variances among pure breeds and for segregation variances between pure breeds at the MQTL. Good ard (1992) proposed a method for genetic evaluation by BLUP using marker and trait information in multibreed populations. His model for means is the same as that presented here, but he uses the Westell method of phantom parents (Westell et al. 1988) to efficiently set up the mixed model equations. His model for variances, however, does not accommodate heterogeneous variances among pure breeds. Further, his formula, for the variance does not have the second term of (14), and thus the segregation variance is not properly modeled (M. E. Goddard, personal communication). Theory presented in this paper does account for heterogeneity of variances among pure breeds and for segregation variances between pure breeds at the MQTL and at the RQTL.

Theory presented here for use of marker information relies on results from Wang et al. (1995). For a purebred population, they gave an exact algorithm to construct \( \mathbf{G}_u \), and to compute its inverse for complete marker data. They also proposed two approximate algorithms to construct \( \mathbf{G}_u \), and to compute its inverse for incomplete marker data. For a multibreed population, the same approximations can be used to construct \( \mathbf{G}_u \), and to compute its inverse for incomplete marker data.

This work was supported in part by the Illinois Agricultural Experiment Station Hatch Project 35-0345 (R.L.F. and T.W.) and 35-0367 (M.G.) and by the United States Department of Agriculture (R.L.F. and T.W.), grant 94-37205-1094.
**LITERATURE CITED**


Lande, R., 1981 The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99: 541–553.


Communicating editor: Z-B. Zeng