Marker assisted selection using best linear unbiased prediction

R.L. Fernando* and M. Grossman

University of Illinois, Department of Animal Sciences, 1207 West Gregory Drive Urbana, IL 61801, USA

(received 12 May 1989; accepted 28 August 1989)

Summary - Best linear unbiased prediction (BLUP) is applied to a mixed linear model with additive effects for alleles at a market quantitative trait locus (MQTL) and additive effects for alleles at the remaining quantitative trait loci (QTL). A recursive algorithm is developed to obtain the covariance matrix of the effects of MQTL alleles. A simple method is presented to obtain its inverse. This approach allows simultaneous evaluation of fixed effects, effects of MQTL alleles, and effects of alleles at the remaining QTLs, using known relationships and phenotypic and marker information. The approach is sufficiently general to accommodate individuals with partial or no marker information. Extension of the approach to BLUP with multiple markers is discussed.

marker-assisted selection - best linear unbiased prediction - genetic marker

INTRODUCTION

Genetic engineering techniques have produced a variety of molecular genetic markers with the potential to identify a large number of genetic polymorphisms (Soller

*Author to whom correspondence should be addressed.
and Beckmann, 1982; Smith and Simpson, 1986; Schumm et al., 1988). Marker-assisted selection is one application of these techniques to animal and plant breeding. Information on marker loci that are linked to quantitative trait loci, together with phenotypic information, could be used to increase genetic progress by increasing accuracy of selection and by reducing generation interval (Soller, 1978; Smith and Simpson, 1986).

Geldermann (1975) proposed a least-squares procedure to estimate effects of marker alleles on quantitative traits. Based on selection index principles, Soller (1978) combined marker information and phenotypic information to obtain genetic evaluations. This method has been used to study the additional genetic progress expected from marker-assisted selection (Soller, 1978; Soller and Beckmann, 1983, Smith and Simpson, 1986). Because of the complex nature of animal breeding data, however, these methods may not be applicable directly to marker-assisted selection with field data.

Data from field-recorded populations are affected by non-genetic nuisance factors, such as age of animal, age of dam, management system, season of birth and herb. Also, non-random mating, selection and overlapping generations contribute to the complexity of the data. Best linear unbiased prediction (BLUP; Henderson, 1973, 1975, 1982) deals with these complications when predicting breeding values from phenotypic data. The objective of this paper is to present methodology for the application of BLUP to marker-assisted selection in animal breeding. Each methodological development is illustrated with a numerical example using a single hypothetical pedigree.

METHODOLOGY

Consider a single polymorphic marker locus (ML), closely linked to a quantitative trait locus (QTL). Let \( M_i^p \) and \( M_i^m \) denote alleles at the ML that individual \( i \) inherited from its paternal (\( p \)) and its maternal (\( m \)) parent, and let \( Q_i^p \) and \( Q_i^m \) denote alleles at the market QTL (MQTL) linked to \( M_i^p \) and \( M_i^m \), as shown below:

\[
\begin{array}{cc}
M_i^p & Q_i^p \\
M_i^m & Q_i^m
\end{array}
\]

Let \( v_i^p \) and \( v_i^m \) be the additive effects of \( Q_i^p \) and \( Q_i^m \). Additive effects of alleles at the remaining QTLs, unlinked to the ML, will be denoted by the residual additive effect \( u_i \). Now, the additive effect for individual \( i \), \( a_i \), can be written as

\[
a_i = v_i^p + v_i^m + u_i
\]

The usual model to obtain BLUP if additive effects, given phenotypic information, is

\[
y_i = x_i' \beta + a_i + e_i
\]

where \( y_i \) is the phenotypic value of individual \( i \), \( x_i' \) is a vector of known constants, \( \beta \) is a vector of unknown fixed effects, and \( e_i \) is a random error. Using eq.(2), BLUP allows information from relatives to contribute to the predictor of \( a_i \) through the
covariance matrix of $a_i$ values. Note that this covariance matrix depends on the

\[ \text{Cov}(a_i | r) = G_{a|r} \]

which is proportional to the numerator relationship matrix (e.g., Henderson, 1976).

When marker information ($m$) is also available, the covariance matrix $a_i$ values is

\[ \text{Cov}(a_i | r, m) = G_{a|r,m} \]

It can be shown that $G_{a|r} \neq G_{a|r,m}$, in general. For example, the covariance between

half-sibs that receive the same ML allele from their common parent is higher than

the covariance between half-sibs that receive different ML alleles. This is because

half-sibs receiving the same ML allele also receive the same MQTL allele with

greater frequency than half-sibs receiving different ML alleles.

A. Marker model

To obtain BLUP with phenotypic and marker information, it is convenient to use

\[ y_i = x_i^T \beta + v_i^p + v_i^m + u_i + e_i \] (3)

which is equivalent to equ.(2). The covariance matrix of $v_i$ values ($G_v$) depends

on relationship and marker information. The covariance matrix of $u_i$ values ($G_u$)

depends only on relationship information and is proportional to the numerator

relationship matrix (e.g., Henderson, 1976). Given the covariance matrices $G_v$ and

$G_u$, BLUPs of $v_i$ and $u_i$ values can be obtained using the mixed model equations

(Henderson, 1973). The inverse of $G_u$, which is required on the mixed model

equations, usually is obtained using an algorithm given by Henderson (1976). A

recursive algorithm to construct $G_v$ is given in section B, and an algorithm to

obtain its inverse is in section C.

B. Covariance matrix of MQTL effects

1. Theory. To construct $G_v$, consider the covariance between additive effects of

MQTL alleles. Without loss of generality, consider only paternal MQTL alleles.

Suppose arbitrary individuals $o$ and $o'$ have sires $s$ and $s'$. The MQTL alleles

inherited by $o$ and $o'$ from their sires are $Q_P$ and $Q_{P'}$, having additive effects $v_P^o$ and

$v_{P'}^{o'}$. For paternal MQTL alleles in $o$ and $o'$, the covariance between their additive

effects $v_P^o$ and $v_{P'}^{o'}$ is

\[ \text{Cov}(v_P^o, v_{P'}^{o'}) = \text{Cov}(v_P^o, v_{P'}^{o'} | Q_P^o \equiv Q_{P'}^{o'}) \cdot P(Q_P^o \equiv Q_{P'}^{o'}) \]

\[ = \text{Var}(v_P^o) \cdot P(Q_P^o \equiv Q_{P'}^{o'}) \] (4)

where $\text{Var}(v_P^o) = \sigma_v^2$ is the additive variance of an MQTL allele and $P(Q_P^o \equiv Q_{P'}^{o'})$

is the probability that $Q_P^o$ is identical by descent to $Q_{P'}^{o'}$. For an arbitrary pair of

individuals, one is not a direct descendent of other. If $o$ is not a direct descendent

of the $o'$, $Q_P^o$ can be identical by descent to $Q_{P'}^{o'}$ in 2 mutually exclusive ways:

1) $Q_P^o$ is identical by descent to the maternal MQTL allele of the sire of $o'$ ($Q_s^{o'}$) and

\[ o' \text{ inherits } Q_s^{o'} \]

2) $Q_P^o$ is identical by descent to the paternal MQTL allele of the sire of $o'$ ($Q_{s'}^{o'}$) and

\[ o' \text{ inherits } Q_{s'}^{o'} \]


If marker information is available, the conditional probability that o' inherits $Q_{s'}^p$, given that o' inherits $M_{s'}^p$, is $(1 - r)$, where $r$ is the recombination rate between the ML and the MQTL. Thus if o' inherits $M_{s'}^p$, the probability in equ.(4) can be calculated recursively as

$$ P(Q_{o'}^e \equiv Q_{o'}^m) = P(Q_{o'}^e \equiv Q_{o'}^p) \cdot (1 - r) + P(Q_{o'}^e \equiv Q_{s'}^m) \cdot r \quad (5) $$

Similarly, if o' inherits $M_{s'}^m$

$$ P(Q_{o'}^e \equiv Q_{o'}^m) = P(Q_{o'}^e \equiv Q_{o'}^p) \cdot r + P(Q_{o'}^e \equiv Q_{s'}^m) \cdot (1 - r) \quad (6) $$

If marker information is not available, so that it is not known whether o' inherits $M_{s'}^p$ or $M_{s'}^m$, 0.5 replaces $r$ in equus.(5) and (6). This is because, in the absence of marker information, $Q_{s'}^p$ and $Q_{s'}^m$ have equal probability of being transmitted to o'.

The above development leads to a tabular method to construct $G_v$, which is similar to the method used to construct the numerator relationship matrix (e.g. Henderson, 1976). Note that $G_v$ has twice as many rows as individuals because each individual has 2 effects: 1 for the paternal and 1 for the maternal MQTL allele. The rows and columns of $G_v$ should be ordered so that the corresponding to progeny follow those for their parents. Let the row indices of $G_v$, corresponding to the effects of MQTL alleles of individual o$(v_{i_s}^p, v_{i_s}^m)$, be $i_{s'}^p, i_{s'}^m$; of its sire $s(v_{i_s}^p, v_{i_s}^m)$, be $i_{s'}^p, i_{s'}^m$; and of its dam $d(v_{i_d}^p, v_{i_d}^m)$, be $i_{d'}^p, i_{d'}^m$. Also, let element $i_j$ of $G_v$ be $g_{ij}$. Then from equus.(4), (5) and (6), the elements of row $i_{o'}^p$, below the diagonal, are obtained as

$$ g_{i_{o'}^p, j} = (1 - \rho_{o'}^p)g_{i_{o'}^p, j} + \rho_{o'}^p g_{i_{o'}^m, j} $$

for $j = 1 \ldots i_{o'}^p - 1$, where $\rho_{o'}^p = r$ if o inherits $M_{s'}^p$ or $\rho_{o'}^p = (1 - r)$ if o inherits $M_{s'}^m$.

Elements of column $i_{o'}^p$, above the diagonal, are obtained from the corresponding row elements because $G_v$ is symmetric. Similarly, elements of row $i_{o'}^m$, below the diagonal, are obtained as

$$ g_{i_{o'}^m, j} = (1 - \rho_{o'}^m)g_{i_{o'}^m, j} + \rho_{o'}^m g_{i_{o'}^p, j} $$

for $j = 1 \ldots i_{o'}^m - 1$, where $\rho_{o'}^m = r$ if o inherits $M_{d'}^p$ and $\rho_{o'}^m = (1 - r)$ if o inherits $M_{d'}^m$.

Elements of column $i_{o'}^m$, above the diagonal, are obtained from the corresponding row elements.

From equ.(4), the diagonal elements of $G_v$ are equal to $\sigma_v^2$. If marker information cannot be used to determine which of the 2 marker alleles o are inherited from its sire or its dam, then 0.5 replaces $\rho_{o'}^p$ in equ.(7a) or $\rho_{o'}^m$ in (7b).

2. Numerical example. Consider the pedigree in Table I. To construct $G_v$, rows and columns are arranged by individual and by paternal and maternal MQTL alleles within individual (Table II). For convenience, we will assume that $\sigma_v^2 = 1$ and that $r = 0.1$. The first two individuals are assumed to be unrelated; thus the upper left $4 \times 4$ submatrix of $G_v$ is the identity matrix. Elements on the diagonal are equal to $\sigma_v^2 = 1$. Now, row elements below the diagonal can be obtained from equus.(7a) and (7b); column elements above the diagonal are obtained by symmetry. Each row element for $v_{i_s}^p$ is equal to $(1 - r) = 0.9$ times the corresponding row element for $v_{i_s}^p$ plus $r = 0.1$ times the corresponding row element for $v_{i_s}^m$. Each row element for $v_{i_d}^m$ is equal to $r = 0.1$ times the corresponding row element for $v_{i_d}^p$ plus $(1 - r) = 0.9$ times the corresponding row element for $v_{i_d}^m$. The ML allele inherited by 4 from
its sire is unknown. Thus, each row element for $v_4^p$ is the mean ($r = 0.5$) of the corresponding row element for $v_1^p$ and for $v_1^m$. Marker information is available for $v_4^m$, so that each row element for $v_4^m$ is $(1 - r) = 0.9$ times the corresponding row element for $v_3^p$ plus $r = 0.1$ times the corresponding row element for $v_3^m$.

Table I. Pedigree with marker information

| Animal | Sire | Dam | Marker inherited from |
|--------|------|-----|----------------------|
|        |      |     |                      |
| 1      | -    | -   | -                    |
| 2      | -    | -   | -                    |
| 3      | 1    | 2   | $M_1^p$              |
| 4      | 1    | 3   | $M_2^m$              |

Table II. Covariance matrix of MQTL effects: $G_v$

|      | $v_1^p$ | $v_1^m$ | $v_2^p$ | $v_2^m$ | $v_3^p$ | $v_3^m$ | $v_4^p$ | $v_4^m$ |
|------|---------|---------|---------|---------|---------|---------|---------|---------|
| $v_1^p$ | 1.0     | 0.0     | 0.0     | 0.0     | 0.9     | 0.0     | 0.5     | 0.81    |
| $v_1^m$ | 0.0     | 1.0     | 0.0     | 0.0     | 0.1     | 0.0     | 0.5     | 0.09    |
| $v_2^p$ | 0.0     | 0.0     | 1.0     | 0.0     | 0.0     | 0.1     | 0.0     | 0.01    |
| $v_2^m$ | 0.0     | 0.0     | 0.0     | 1.0     | 0.0     | 0.9     | 0.0     | 0.09    |
| $v_3^p$ | 0.9     | 0.1     | 0.0     | 0.0     | 1.0     | 0.0     | 0.5     | 0.9     |
| $v_3^m$ | 0.0     | 0.0     | 0.1     | 0.9     | 0.0     | 1.0     | 0.0     | 0.1     |
| $v_4^p$ | 0.5     | 0.5     | 0.0     | 0.0     | 0.5     | 0.0     | 1.0     | 0.45    |
| $v_4^m$ | 0.81    | 0.09    | 0.01    | 0.09    | 0.9     | 0.1     | 0.45    | 1.0     |

C. Algorithm for inverting $G_v$

1. Theory. The approach taken here follows that by Quaas et al. (1984) and Quaas (1988) to invert the matrix of additive relationships. We define a linear model to relate the effect of the paternal MQTL allele of an individual ($o$) to effects of paternal and maternal MQTL alleles of its sire ($s$)

$$v_o^p = (1 - \rho_o^p) v_s^p + \rho_o^p v_s^m + \epsilon_o^p$$  \hspace{1cm} (8a)

where $\epsilon_o^p$ is a residual effect. Similarly, a linear model for effect of the maternal MQTL allele of $o$ is

$$v_o^m = (1 - \rho_o^m) v_d^p + \rho_o^m v_d^m + \epsilon_o^m$$  \hspace{1cm} (8b)

It can be shown that the residuals $\epsilon_o^p$ in equ.(8a) and $\epsilon_o^m$ in (8b) have a diagonal covariance matrix ($G_\epsilon$; see Appendix). Now, the vector of effects of MQTL alleles ($v$) can be written as

$$v = P v + \epsilon$$  \hspace{1cm} (9)
where $P$ is a matrix with each row containing only two non-zero elements, if the parent is known or containing only zeros, if the parent is unknown; and where $e$ is a vector of residuals. For example, row $i^p$ will have $(1 - \rho^p)$ in column $i^p$ and $\rho^p$ in column $i^m$, if the sire of $i$ is known. Similarly, row $i^m$ will have $(1 - \rho^m)$ in the column $i^d$ and $\rho^m$ in column $i^m$, if the dam of $i$ is known.

To proceed, we need the diagonal elements of $G_e$. Consider, for example, the variance of $\varepsilon^p$. From equ.(8a), if the sire of $o$ is known

$$\text{Var}(\varepsilon^p) = (1 - \rho^p)^2 \cdot \text{Var}(\varepsilon^p) + (\rho^p)^2 \cdot \text{Var}(\varepsilon^m) + 2(1 - \rho^p)\rho^p \cdot \text{Cov}(\varepsilon^p, \varepsilon^m)$$

because effects of MQTL alleles of sire $s$ are uncorrelated with residuals of its offspring $o$ (see Appendix). Hence

$$\text{Var}(\varepsilon^p) = \text{Var}(\varepsilon^p) - (1 - \rho^p)^2 \cdot \text{Var}(\varepsilon^p) - (\rho^p)^2 \cdot \text{Var}(\varepsilon^m) - 2(1 - \rho^p)\rho^p \cdot \text{Cov}(\varepsilon^p, \varepsilon^m)$$

$$= 2\sigma^2(1 - \rho^p)\rho^p (1 - F_s)$$

where $F_s$ is the inbreeding of sire $s$. Now, equ.(10) can be written as

$$\text{Var}(\varepsilon^m) = 2\sigma^2(1 - \rho^m)\rho^m (1 - F_d)$$

where $(1 - \rho^m)\rho^m = (1 - r)r$ for $\rho^m = r$ or for $\rho^m = (1 - r)$. When the sire is not inbred: $\text{Var}(\varepsilon^p) = 2\sigma^2(1 - r)r$, if marker information is available; or $\text{Var}(\varepsilon^p) = \sigma^2/2$, if marker information is not available. If the sire is not known, $\text{Var}(\varepsilon^p) = \sigma^2/2$.

Similarly, if dam of $o$ is known, the variance of $e^m$ is

$$\text{Var}(\varepsilon^m) = 2\sigma^2(1 - \rho^m)\rho^m (1 - F_d)$$

where $(1 - \rho^m)\rho^m = (1 - r)r$ for $\rho^m = r$ or for $\rho^m = (1 - r)$ and where $F_d$ is the inbreeding of dam $d$. When the dam is not inbred: $\text{Var}(\varepsilon^m) = 2\sigma^2(1 - r)r$, if marker information is available; or $\text{Var}(\varepsilon^m) = \sigma^2/2$, if marker information is not available. If the dam is not known, $\text{Var}(\varepsilon^m) = \sigma^2/2$.

Rearranging (9), $v$ can be written as

$$v = (I - P)^{-1}e$$

for non-singular $(I - P)$, and thus $G_v$ can be written as

$$G_v = (I - P)^{-1}G_e(I - P')^{-1}$$

From equ.(14), it is clear that $G_v^{-1}$ can be written as

$$G_v^{-1} = (I - P')G_e^{-1}(I - P)$$

As shown earlier, $P$ has a simple structure, with each row containing at most 2 non-zero elements, and $G_e^{-1}$ is diagonal.

To obtain the rules for inverting $G_v^{-1}$, equ.(15) is written as

$$G_v^{-1} = QG_e^{-1}Q'$$

where $Q = (I - P')$. Because $G_e$ is diagonal, equ.(16) can be written as
where $n$ is number of individuals in the pedigree, $q_j$ is column $j$ of $Q$, and $d_j$ is diagonal element $j$ of $G^{-1}$. By definition of $Q$, element $j$ of $q_j$ is unity. Further, $q_j$ will have, at most, only 2 other non-zero elements; for $j = i_o^p$, element $i_o^p$ equals $-(1 - p_o^p)$ and element $i_m^p$ equals $-p_o^m$, if the sire of $o$ is known. Similarly, for $j = i_o^m$, element $i_o^m$ equals $-(1 - p_o^m)$ and element $i_m^m$ equals $-p_o^m$, if the dam of $o$ is known. Thus, given parent and marker information of an individual, the contributions to $G^{-1}$, corresponding to effects of paternal and maternal MQTL alleles of the individual, are easily obtained.

Now, to obtain the inverse of $G$: 1) calculate diagonals of $G$: when the parent is known, the diagonal is given by eqns. (12a) or (12b), and when the parent is unknown, the diagonal is $\sigma_v^2$; 2) set $G^{-1}$ to the null matrix; 3) for each offspring $o$, with sire $s$ and dam $d$, add the following to the indicated elements of $G^{-1}$:

\begin{align*}
\text{if sire is known, add } & (1 - \rho_o^p)^2 d_{i_o^p} \text{ to diagonal element } i_o^p, i_o^p; \\
& -(1 - \rho_o^p) d_{i_o^p} \text{ to elements } i_s^p, i_o^m \text{ and } i_o^m, i_s^m; \\
& (\rho_o^p)^2 d_{i_o^m} \text{ to diagonal element } i_m^p, i_m^m; \\
& -\rho_o^m d_{i_o^m} \text{ to elements } i_s^m, i_o^m \text{ and } i_o^m, i_s^m;
\end{align*}

if dam is known, add $(1 - \rho_o^m)^2 d_{i_o^m}$ to diagonal element $i_o^d, i_o^d$;

\begin{align*}
& (1 - \rho_o^m) \rho_o^m d_{i_o^m} \text{ to elements } i_o^d, i_d^m \text{ and } i_d^m, i_o^d; \\
& (1 - \rho_o^m) d_{i_o^m} \text{ to elements } i_d^m, i_o^m \text{ and } i_o^m, i_d^m; \\
& (\rho_o^m)^2 d_{i_o^m} \text{ to diagonal element } i_d^m, i_d^m; \\
& -\rho_o^m d_{i_o^m} \text{ to elements } i_o^m, i_o^m \text{ and } i_o^m, i_d^m; \\
\end{align*}

and always add $d_{i_o^m}$ to element $i_o^m, i_o^m$ and $d_{i_o^m}$ to element $i_o^m, i_o^m$.

2. Numerical example. Consider the pedigree in Table I. To construct $G$, we again take $\sigma_v^2 = 1$ and $r = 0.1$. Because the parents of individuals 1 and 2 are not known, the first 4 elements on the diagonal of $G$ are $\sigma_v^2 = 1$. For individual 3, each parent is known and marker information is available. Thus, from eqns.(12a) and (12b), the two diagonals of $G$ corresponding to effects of paternal and maternal MQTL alleles of individual 3 are $2(1 - r)r = 0.18$. Each parent of individual 4 is also known, but the marker inherited from the sire is not known. Therefore, the diagonal of $G$ corresponding to $v_4^p$ is 0.5, and that corresponding to $v_4^m$ is $2(1 - r)r = 0.18$.

The $P$ matrix for this example is given in Table III. The first 4 rows of $P$ are null because parents of the first 2 individuals are not known. The sire of individual 3 is 1, and $M_3^p$ was transmitted to 3. Thus, the row corresponding to $v_3^p$ has $(1 - r) = 0.9$ in the column corresponding to $v_3^p$ and $r = 0.1$ in the column corresponding to $v_3^m$. Similarly, the dam of individual 3 is 2, and $M_3^m$ was transmitted to 3. Thus, the row corresponding to $v_3^m$ has $r = 0.1$ in the column corresponding to $v_3^p$ and $(1 - r) = 0.9$ in the column corresponding to $v_3^m$. The sire of individual 4 is 1, but marker information is not available. Thus, the row corresponding to $v_4^p$ has 0.5 in the columns corresponding to $v_4^p$ and $v_4^m$. The dam of individual 4 is 3, and $M_4^p$...
was transmitted to 4. Thus, the row corresponding to \( v^m_4 \) has \((1 - r) = 0.9\) in the column corresponding to \( v^p_3 \) and \( r = 0.1\) in the column corresponding to \( v^m_3 \).

The matrix \( Q = (I - P') \) is given in Table IV. The product \( QG^{-1}Q' \) is given in Table V. It can be verified that this is identical to the inverse of the matrix \( G_v \) in Table II.

**Table III. \( P \) matrix**

| \( v^p_1 \) | \( v^m_1 \) | \( v^p_2 \) | \( v^m_2 \) | \( v^p_3 \) | \( v^m_3 \) | \( v^p_4 \) | \( v^m_4 \) |
|------|------|------|------|------|------|------|------|
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.9  | 0.1  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.5  | 0.5  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.9  | 0.1  | 0.0  | 0.0  |

**Table IV. \( Q \) matrix**

| \( v^p_1 \) | \( v^m_1 \) | \( v^p_2 \) | \( v^m_2 \) | \( v^p_3 \) | \( v^m_3 \) | \( v^p_4 \) | \( v^m_4 \) |
|------|------|------|------|------|------|------|------|
| 1.0  | 0.0  | 0.0  | 0.0  | -0.9 | 0.0  | -0.5 | 0.0  |
| 0.0  | 1.0  | 0.0  | 0.0  | -0.1 | 0.0  | -0.5 | 0.0  |
| 0.0  | 0.0  | 1.0  | 0.0  | 0.0  | -0.1 | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 1.0  | 0.0  | -0.9 | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 1.0  | 0.0  | 0.0  | -0.9 |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 1.0  | 0.0  | -0.1 |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 1.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 1.0  |

**Table V. \( QG^{-1}Q' \)**

| \( v^p_1 \) | \( v^m_1 \) | \( v^p_2 \) | \( v^m_2 \) | \( v^p_3 \) | \( v^m_3 \) | \( v^p_4 \) | \( v^m_4 \) |
|------|------|------|------|------|------|------|------|
| 6.0  | 1.0  | 0.0  | 0.0  | -5.0 | 0.0  | -1.0 | 0.0  |
| 1.0  | 1.556| 0.0  | 0.0  | -0.556| 0.0  | -1.0 | 0.0  |
| 0.0  | 0.0  | 1.056| 0.5  | 0.0  | -0.556| 0.0  | 0.0  |
| 0.0  | 0.0  | 0.5  | 5.5  | 0.0  | -5.0  | 0.0  | 0.0  |
| -5.0 | -0.556| 0.0  | 0.0  | 10.056| 0.5  | 0.0  | -5.0 |
| 0.0  | 0.0  | -0.556| -5.0 | 0.5  | 5.611 | 0.0  | -0.556|
| -1.0 | -1.0 | 0.0  | 0.0  | 0.0  | 2.0  | 0.0  | 0.0  |
| 0.0  | 0.0  | 0.0  | 0.0  | -5.0 | -0.556| 0.0  | 5.556|
**D. BLUP with multiple markers**

If information on another marker locus linked to a QTL is available, the model can be expanded to include effects of alleles of this MQTL. This approach, however, results in \(2n\) additional equations for each marker introduced into the analysis. Thus, for a large number of individuals \(n\) and a large number of MQTLs, solving the mixed model equations may not be feasible. An alternative would be to use equ.(2), with

\[
a_i = \sum_k (v_{ki}^p + v_{ki}^m) + u_i
\]

where \(v_{ki}^p\) and \(v_{ki}^m\) are effects of paternal and maternal alleles of the \(k\)th MQTL. The covariance matrix of effects of MQTL alleles at each locus \((G_{v_k})\) can be constructed using the tabular method described in Section II.B. Then, assuming gametic equilibrium, the covariance of matrix \(a_i\) values \((G_{a|r,m})\) can be obtained as

\[
G_{a|r,m} = Z(\sum_k G_{v_k})Z' + G_u
\]

where \(Z\) is a \(n \times 2n\) matrix with elements for row \(i\) containing a 1 corresponding to each of the paternal and maternal MQTL effects of individual \(i\) and zeros for the remaining elements. The problem with this approach, however, is that it could not be applied to large systems, unless a simple algorithm to invert \(G_{a|r,m}\) is available.

**DISCUSSION**

Results presented here are an application of BLUP to marker-assisted selection. This is a generalization of the method presented by Soller (1978) and Soller and Beckmann (1983). This generalization allows simultaneous evaluation of fixed effects, MQTL effects and the residual QTL effects, using known relationships and phenotypic and marker information. It is sufficiently general to accommodate individuals with partial or no marker information.

Several authors have calculated the additional genetic progress expected from marker-assisted selection (Soller, 1978, Soller and Beckmann, 1983; Smith and Simpson, 1986). Because the method presented here is a generalization of the method considered by these authors, their results give an indication of the advantage expected by using marker-assisted BLUP.

Application of this procedure requires knowledge of the recombination rate \((r)\) between the marker and the MQTL and the variance of the additive effect of the MQTL alleles \((\sigma_r^2)\). Assuming that effects of MQTL alleles are normally distributed, the model presented here could be used to estimate \(r\) and \(\sigma_r^2\) by restricted maximum likelihood (REML; Patterson and Thompson, 1971). The robustness of REML estimation, with respect to the distribution of effects of MQTL alleles, needs to be examined.
REFERENCES

Geldermans H. (1975) Investigation on inheritance of quantitative characters in animals by gene markers. 1. Methods. Theor. Appl. Genet. 46, 319-330

Henderson C.R. (1973) Sire evaluation and genetic trends. In: Animal Breeding and Genetics Symposium in Honor of Dr Jay L. Lush, American Society of Animal Science and American Dairy Science Association, Champaign, IL, pp. 10-41

Henderson C.R. (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423-439

Henderson C.R. (1976) A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32, 69-83

Henderson C.R. (1982) Best linear unbiased prediction in populations that have undergone selection. In: World Congress on Sheep and Beef Cattle Breeding, (Barton R.A. & Smith W.C. ed.) vol. 1, Dunsmore Press, Palmerston North, NZ, pp. 191-200

Henderson C.R. (1988) Use of an average numerator relationship matrix for multiple-sire joining. J. Anim. Sci. 66, 1614-1621

Patterson H.D. & Thompson R. (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554

Quaas R.L. (1988) Additive genetic model with groups and relationships. J. Dairy Sci. 71, 1338-1345

Quaas R.L., Anderson R.D. & Gilmour A.R. (1984) BLUP School Handbook; Use of Mixed Models for Prediction and for Estimation of (Co)variance Components. Animal Genetics and Breeding Unit, University of New England, N.S.W. 2351, Australia.

Schumm J.W., Knowlton R.G., Braman J.C., Barker D.F., Botstein D., Akots G., Brown V.A., Gravius T.C., Helms C., Hsiao K., Rediker K., Thurston J.G. & Donis-Keller H. (1988) Identification of more than 500 RFLPs by screening random genomic clones. Am. J. Hum. Genet. 42, 143-159

Smith C. & Simpson S.P. (1986) The use of genetic polymorphisms in livestock improvement. J. Anim. Breeding. Genet. 103, 205-217

Soller M. (1978) The use of loci associated with quantitative traits in dairy cattle improvement. Anim. Prod. 27, 133-139

Soller M., Beckmann J.S. (1982) Restriction fragment length polymorphisms and genetic improvement. In: World Congress on Genetics Applied to Livestock Production, Madrid, vol. 6, pp. 396-404

Soller M., Beckmann J.S. (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. 67, 25-33
Proof that $G_{\xi}$ is diagonal

Let $o$ be an individual that is not a direct descendant of $o'$. From equ.(8a) and (8b), the additive effects of the MQTL alleles of $o$ and $o'$ are

$$v_o^z = (1 - \rho_o^z)v_{\xi}^p + \rho_o^z v_{\xi}^m + \epsilon_o^z$$  \hspace{1cm} (A1)

and

$$v_{o'}^z = (1 - \rho_{o'}^z)v_{\xi}^p + \rho_{o'}^z v_{\xi}^m + \epsilon_{o'}^z$$  \hspace{1cm} (A2)

where $z$ can take values $p$ or $m$, $\xi = s$ when $z = p$, or $\xi = d$ when $z = m$. Similarly, $z'$ can take values $p$ or $m$, $\xi' = s'$ when $z' = p$, or $\xi' = d$ when $z' = m$. Note that for an arbitrary pair of individuals, one is not direct descendant of the other. Therefore, to prove that $G_{\xi}$ is diagonal, it is sufficient to show that the covariance between $\epsilon_o^z$ and $\epsilon_{o'}^{z'}$ is null.

From equ.(A1) and (A2), the covariance between additive effects of MQTL alleles $v_o^z$ and $v_{o'}^{z'}$ can be written as

$$\text{Cov}(v_o^z, v_{o'}^{z'}) = \text{Cov}(v_o^z, (1 - \rho_{o'}^{z'})v_{\xi}^p + \rho_{o'}^{z'} v_{\xi}^m + \epsilon_{o'}^{z'})$$

$$= (1 - \rho_{o'}^{z'})\text{Cov}(v_o^z, v_{\xi}^p) + \rho_{o'}^{z'}\text{Cov}(v_o^z, v_{\xi}^m)$$

$$+ \text{Cov}(v_o^z, \epsilon_{o'}^{z'})$$  \hspace{1cm} (A3)

But, from equ.(7a) and (7b)

$$\text{Cov}(v_o^z, v_{o'}^{z'}) = (1 - \rho_{o'}^{z'})\text{Cov}(v_o^z, v_{\xi}^p) + \rho_{o'}^{z'}\text{Cov}(v_o^z, v_{\xi}^m)$$  \hspace{1cm} (A4)

Thus, for equ.(A3) to equal (A4), the third term in equ.(A3), $\text{Cov}(v_o^z, \epsilon_{o'}^{z'})$, must be zero. The same reasoning can be used to show that $\text{Cov}(v_{\xi}^p, \epsilon_{o'}^{z'})$ and $\text{Cov}(v_{\xi}^m, \epsilon_{o'}^{z'})$ are zero. Therefore, given that $\text{Cov}(v_o^z, \epsilon_{o'}^{z'})$, $\text{Cov}(v_{\xi}^p, \epsilon_{o'}^{z'})$ and $\text{Cov}(v_{\xi}^m, \epsilon_{o'}^{z'})$ are zero, $\text{Cov}(\epsilon_o^z, \epsilon_{o'}^{z'})$ must be zero.

Further, taking $o$ to be the a parent of $o'$, the residual ($\epsilon_{o'}^{z'}$) in equ.(A2) is uncorrelated with $v_{\xi}^p$ and with $v_{\xi}^m$ in equ.(A2), because $\text{Cov}(v_o^z, \epsilon_{o'}^{z'}) = 0$, as shown above. The result that be effect of each MQTL allele of a parent is uncorrelated with the residual ($\epsilon_{o'}^{z'}$) of its offspring was used to obtain equ.(10)