465

Original article

A reduced animal model with elimination of quantitative trait loci equations for marker-assisted selection

S Saito [1], H Iwaisaki [1,2]*

[1] Graduate School of Science and Technology;
[2] Department of Animal Science, Faculty of Agriculture,
Niigata University, Niigata 950-21, Japan

Summary – A reduced animal model (RAM) version of the method with the animal model proposed by Hoeschele for marker-assisted selection is presented. The current RAM approach allows simultaneous evaluation of fixed effects, the total additive genetic merits for parent animals and the additive effects due to quantitative trait loci linked to the marker locus only for animals which have the marker data or provide relationship ties among descendant animals with known marker data. An appropriate covariance matrix of the residual effects is given, and formulae for backsolving for non-parent animals are presented. A numerical example is also given.

marker-assisted selection / best linear unbiased prediction / reduced animal model / total additive genetic effect / additive effect of marked QTL alleles

Résumé – *Un modèle animal réduit avec élimination d'équations relatives aux locus de caractères quantitatifs pour la sélection assistée par marqueurs. Le modèle animal proposé par I Hoeschele pour la sélection assistée par marqueurs est modifié ici en modèle animal réduit (MAR). Cette approche MAR permet d'évaluer simultanément les effets fixes, les valeurs génétiques additives des individus parents et les effets additifs de locus liés aux locus marqueurs pour les seuls individus marqués ou qui fournissent des liens de parenté entre des descendants marqués. La matrice de covariance résiduelle correspondante est donnée, ainsi que les formules permettant de remonter aux individus non parents. Un exemple numérique est également traité.*

sélection assistée par marqueurs / meilleure prédiction linéaire sans biais / modèle animal réduit / valeur génétique additive totale / effet additif de locus quantitatif marqué

---

* Correspondence and reprints

INTRODUCTION

A procedure for marker-assisted selection (MAS) using best linear unbiased prediction (BLUP; Henderson, 1973, 1975, 1984) was first proposed by Fernando and Grossman (1989), showing how marker information can be utilized in an animal model (AM) for simultaneous evaluation of fixed effects, additive genetic effects due to quantitative trait loci (QTL) unlinked to the marker loci (ML) and additive effects due to marked QTL (MQTL). Later, certain authors presented various types of procedures to incorporate marker information in BLUP, taking into consideration multiple markers, using a reduced animal model (RAM), or combining the MQTL effects and the effects of alleles at the remaining QTL into the total additive genetic merits.

Goddard (1992) extended Fernando and Grossman's (1989) model for flanking markers and discussed the use of RAM. Cantet and Smith (1991) derived a RAM version of the AM model of Fernando and Grossman (1989). Also, an AM method to reduce the number of equations per animal to one was presented by van Arendonk et al (1994), combining information on MQTL and QTL unlinked to ML into one numerator relationship matrix. A RAM approach to the AM of van Arendonk et al (1994) is also available (Saito and Iwaisaki, 1997).

Hoeschele (1993) worked with an AM of the total additive genetic merits and the additive effects for MQTL alleles, and indicated that if some of the animals to be evaluated do not have marker data and do not provide relationship ties among genotyped descendants with known marker data, the MQTL equations for such animals can be eliminated, showing that the inverse of a covariance matrix among the total additive genetic merits and the needed additive effects of the MQTL alleles can be obtained directly. When applied to genetic evaluations in which only a small fraction of the animals are genotyped for ML and the remaining fraction do not provide marker data, Hoeschele's (1993) procedure can have the large advantage of reducing the number of equations to be solved.

In this paper, a RAM version of the model of Hoeschele (1993) is described. The current approach does not require the MQTL equations for parent animals that were not marker genotyped and that do not provide relationship ties among marker genotyped descendants.

THEORY

For simplicity, one MQTL is assumed in the derivations.

A RAM for MAS

If each animal in the relevant population has only one observation, the AM of Fernando and Grossman (1989) can be arranged, and a RAM can be obtained as described by Cantet and Smith (1991).

The AM is written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{Z}\mathbf{P}\mathbf{v} + \mathbf{e} \qquad [1]$$

where $\mathbf{y}_{(n\times1)}$ is the vector of observations, $\boldsymbol{\beta}_{(f\times1)}$ is the vector of fixed effects, $\mathbf{u}_{(q\times1)}$ is the random vector of the additive genetic effects due to QTL not linked to the ML, $\mathbf{v}_{(2q\times1)}$ is the random vector of the additive effects of the MQTL alleles, $\mathbf{e}_{(n\times1)}$ is the random vector of residual effects, and $\mathbf{X}_{(n\times f)}, \mathbf{Z}_{(n\times q)}$ and $\mathbf{P}_{(q\times2q)}$ are the known incidence matrices, respectively. The subscripts represent the sizes of the vectors and the matrices. The expectation and dispersion matrices for the random effects are usually assumed as

$$
\mathrm{E}\begin{bmatrix}\mathbf{u}\\\mathbf{v}\\\mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\end{bmatrix}\qquad \mathrm{Var}\begin{bmatrix}\mathbf{u}\\\mathbf{v}\\\mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{A}_u\sigma_u^2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_v\sigma_v^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2\end{bmatrix}
$$

where $\mathbf{A}_u$ is the numerator relationship matrix for the QTL unlinked to the ML, $\mathbf{A}_v$ is the gametic relationship matrix for the MQTL, $\mathbf{I}$ is an identity matrix, and σ_u^2, σ_v^2 and σ_e^2 are the variance components for the polygenic effects due to QTL unlinked to the ML, the additive effects of MQTL alleles and the residual effects, respectively. The mixed-model equations (MMEs) are given as

$$
\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{ZP} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z}+\mathbf{A}_u^{-1}\alpha_u & \mathbf{Z}'\mathbf{ZP} \\ \mathbf{P}'\mathbf{Z}'\mathbf{X} & \mathbf{P}'\mathbf{Z}'\mathbf{Z} & \mathbf{P}'\mathbf{Z}'\mathbf{ZP}+\mathbf{A}_v^{-1}\alpha_v \end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}^o\\\widehat{\mathbf{u}}\\\widehat{\mathbf{v}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{y}\\\mathbf{Z}'\mathbf{y}\\\mathbf{P}'\mathbf{Z}'\mathbf{y}\end{bmatrix} \qquad [2]
$$

where $\alpha_u = \sigma_e^2/\sigma_u^2$ and $\alpha_v = \sigma_e^2/\sigma_v^2$.

Then the vectors $\mathbf{y}$, $\mathbf{u}$ and $\mathbf{v}$ in equation [1] can be partitioned as

$$
\mathbf{y}=\begin{bmatrix}\mathbf{y}_p' \ \mathbf{y}_o'\end{bmatrix}',\quad \mathbf{u}=\begin{bmatrix}\mathbf{u}_p' \ \mathbf{u}_o'\end{bmatrix}' \text{ and } \mathbf{v}=\begin{bmatrix}\mathbf{v}_p' \ \mathbf{v}_o'\end{bmatrix}'
$$

respectively, where the subscripts p and o refer to animals with progeny and without progeny, respectively. Also, $\mathbf{u}_o$ and $\mathbf{v}_o$ are further expressed as follows

$$
\mathbf{u}_o = \mathbf{T}\mathbf{u}_p + \mathbf{m} \qquad [3]
$$

and

$$
\mathbf{v}_o = \mathbf{B}\mathbf{v}_p + \boldsymbol{\varepsilon} \qquad [4]
$$

where $\mathbf{T}$ is a matrix relating $\mathbf{u}_o$ to $\mathbf{u}_p$ and has zeros except for 0.5 in the column pertaining to a known parent, $\mathbf{m}$ is a vector of the Mendelian sampling effects, $\mathbf{B}$ is a matrix relating $\mathbf{v}_o$ to $\mathbf{v}_p$ and contains at most four non-zero elements in each row if the parental origin of marker alleles cannot be determined (Hoeschele, 1993; van Arendonk et al, 1994; Wang et al, 1995), and $\boldsymbol{\varepsilon}$ is a vector of segregation residuals. Thus equation [1] can be written as a RAM by

$$
\begin{bmatrix}\mathbf{y}_p\\\mathbf{y}_o\end{bmatrix}=\begin{bmatrix}\mathbf{X}_p\\\mathbf{X}_o\end{bmatrix}\boldsymbol{\beta}+\begin{bmatrix}\mathbf{Z}_p\\\mathbf{Z}_o\mathbf{T}\end{bmatrix}\mathbf{u}_p+\begin{bmatrix}\mathbf{Z}_p\mathbf{P}_p\\\mathbf{Z}_o\mathbf{P}_o\mathbf{B}\end{bmatrix}\mathbf{v}_p+\begin{bmatrix}\mathbf{e}_p\\\mathbf{Z}_o\mathbf{m}+\mathbf{Z}_o\mathbf{P}_o\boldsymbol{\varepsilon}+\mathbf{e}_o\end{bmatrix} \qquad [5]
$$

and equation [5] with $\mathbf{Z}_o = \mathbf{I}_o$, where $\mathbf{I}_o$ is an identity matrix, can be rewritten as

$$
\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_t\mathbf{u}_p + \mathbf{W}\mathbf{v}_p + \boldsymbol{\phi} \qquad [6]
$$

using the appropriate matrices $\mathbf{Z_t}$ and $\mathbf{W}$.

With this RAM, the assumptions for expectation and dispersion parameters of $\mathbf{u}_p$, $\mathbf{v}_p$ and $\boldsymbol{\phi}$ are

$$\mathrm{E}\begin{bmatrix}\mathbf{u}_p \\ \mathbf{v}_p \\ \boldsymbol{\phi}\end{bmatrix} = \begin{bmatrix}0 \\ 0 \\ 0\end{bmatrix} \qquad \mathrm{Var}\begin{bmatrix}\mathbf{u}_p \\ \mathbf{v}_p \\ \boldsymbol{\phi}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{u_p}\sigma_u^2 & 0 & 0 \\ 0 & \mathbf{A}_{v_p}\sigma_v^2 & 0 \\ 0 & 0 & \mathbf{R}\end{bmatrix}$$

where the matrices $\mathbf{A}_{u_p}$ and $\mathbf{A}_{v_p}$ are appropriate submatrices of $\mathbf{A}_u$ and $\mathbf{A}_v$ respectively, and

$$\mathbf{R} = \begin{bmatrix}\mathbf{I}_p\sigma_e^2 & 0 \\ 0 & \mathbf{D}\sigma_u^2 + \mathbf{P}_o\mathbf{G}_\varepsilon\mathbf{P}_o'\sigma_v^2 + \mathbf{I}_o\sigma_e^2\end{bmatrix} \qquad [7]$$

where $\mathbf{I}_p$ is an identity matrix, $\mathbf{D}$ is a diagonal matrix with diagonal elements, $d_i = 1 - 0.25(\delta_{ss(i)} + \delta_{dd(i)})$, where $\delta_{ss(i)}$ and $\delta_{dd(i)}$ are the diagonal elements of $\mathbf{A}_{u_p}$ corresponding to the sire and the dam of animal i, and $\mathbf{G}_\varepsilon$ is a block diagonal matrix (Saito and Iwaisaki, 1996) in which each block is calculated as

$$\mathbf{G}_{\varepsilon(i)} = \begin{bmatrix}1 & f_i \\ f_i & 1\end{bmatrix} - \mathbf{B}_i\mathbf{A}_{v_{p(i)}}\mathbf{B}_i' \qquad [8]$$

where f_i is the conditional inbreeding coefficient of animal i for the MQTL, given the marker information, according to Wang et al (1995).

Consequently, the MMEs for equation [6] are given by

$$\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z_t} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{W} \\ \mathbf{Z}_t'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_t'\mathbf{R}^{-1}\mathbf{Z_t} + \mathbf{A}_{u_p}^{-1}\sigma_u^{-2} & \mathbf{Z}_t'\mathbf{R}^{-1}\mathbf{W} \\ \mathbf{W}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{Z_t} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{A}_{v_p}^{-1}\sigma_v^{-2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}^o \\ \widehat{\mathbf{u}}_p \\ \widehat{\mathbf{v}}_p\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_t'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{W}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

$$[9]$$

The information on recombination rate between the ML and the MQTL and the variance components required in the BLUP approaches for MAS could be obtained, for instance, by the restricted maximum likelihood or the maximum likelihood procedures (eg, Weller and Fernando, 1991; van Arendonk et al, 1993; Grignola et al, 1994).

The RAM containing the total additive genetic effects and only the MQTL effects needed

Consider the following transformation matrix

$$\mathbf{H} = \begin{bmatrix}\mathbf{I}_{(q_p)} & \mathbf{P}_p \\ 0 & \mathbf{I}_{(2q_p)}\end{bmatrix} \qquad [10]$$

with

$$\mathbf{H}\begin{bmatrix}\mathbf{u}_p \\ \mathbf{v}_p\end{bmatrix} = \begin{bmatrix}\mathbf{a}_p \\ \mathbf{v}_p\end{bmatrix}$$

where $\mathbf{I}_{(q_p)}$ and $\mathbf{I}_{(2q_p)}$ are identity matrices, $\mathbf{P}_p$ is the $q_p \times 2q_p$ incidence matrix, and $\mathbf{a}_p$ is the $q_p \times 1$ subvector of $\mathbf{a}$, the vector of the overall genetic values.

Then equation [6] can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{Z_t} \ \mathbf{W}) \begin{pmatrix} \mathbf{u}_p \\ \mathbf{v}_p \end{pmatrix} + \boldsymbol{\phi} \qquad [11]$$

and using equation [10] in equation [11] gives

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{Z_t}\mathbf{W})\mathbf{H}^{-1}\mathbf{H} \begin{pmatrix} \mathbf{u}_p \\ \mathbf{v}_p \end{pmatrix} + \boldsymbol{\phi}$$
$$= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z_t}\mathbf{a}_p + \mathbf{L}\mathbf{v}_p + \boldsymbol{\phi} \qquad [12]$$

where $\mathbf{L} = -\mathbf{Z_t}\mathbf{P}_p + \mathbf{W}$. Also, the inverse covariance matrix of the total additive genetic effects and the additive effects of MQTL alleles is given by $(\mathbf{H}')^{-1}[\mathrm{Var}(\mathbf{u}_p' \ \mathbf{v}_p')']^{-1}\mathbf{H}^{-1}$ (Hoeschele, 1993), or

$$\mathrm{Var}\begin{bmatrix} \mathbf{a}_p \\ \mathbf{v}_p \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{u_p}^{-1}\sigma_u^{-2} & -\mathbf{A}_{u_p}^{-1}\mathbf{P}_p\sigma_u^{-2} \\ -\mathbf{P}_p'\mathbf{A}_{u_p}^{-1}\sigma_u^{-2} & \mathbf{P}_p'\mathbf{A}_{u_p}^{-1}\mathbf{P}_p\sigma_u^{-2} + \mathbf{A}_{v_p}^{-1}\sigma_v^{-2} \end{bmatrix} = \mathbf{G}_p^{-1} \qquad [13]$$

Therefore, the expectation and dispersion parameters of $\mathbf{a}_p$, $\mathbf{v}_p$ and $\boldsymbol{\phi}$ in equation [12] are given as

$$\mathrm{E}\begin{bmatrix} \mathbf{a}_p \\ \mathbf{v}_p \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \qquad \mathrm{Var}\begin{bmatrix} \begin{bmatrix} \mathbf{a}_p \\ \mathbf{v}_p \end{bmatrix} \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{G}_p \end{bmatrix} & \mathbf{0} \\ & & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}$$

Then the MMEs for equation [12] are

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z_t} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{L} \\ \mathbf{Z_t}'\mathbf{R}^{-1}\mathbf{X} & \begin{bmatrix} \mathbf{Z_t}'\mathbf{R}^{-1}\mathbf{Z_t} & \mathbf{Z_t}'\mathbf{R}^{-1}\mathbf{L} \\ \mathbf{L}'\mathbf{R}^{-1}\mathbf{Z_t} & \mathbf{L}'\mathbf{R}^{-1}\mathbf{L} \end{bmatrix} + \mathbf{G}_p^{-1} \\ \mathbf{L}'\mathbf{R}^{-1}\mathbf{X} & \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}^o \\ \widehat{\mathbf{a}}_p \\ \widehat{\mathbf{v}}_p \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z_t}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{L}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \qquad [14]$$

Now, for animals which have unknown marker data and have only one progeny with known marker data in the relevant population, the equations for the additive effects of MQTL alleles for BLUP may not be needed and may be eliminated (Hoeschele, 1993). If some of the additive effects of the MQTL alleles with a non-parent animal, its sire or its dam can be eliminated, then formula [7] is no longer true.

Since the vectors of the total additive genetic effects and the needed additive effects of MQTL alleles can be represented as $\mathbf{g}_p = [\mathbf{a}_p' \ \mathbf{v}_p^{*'}]'$ for parent animals and as $\mathbf{g}_o = [\mathbf{a}_o' \ \mathbf{v}_o^{*'}]'$ for non-parent animals, where $\mathbf{v}_p^*$ and $\mathbf{v}_o^*$ are subvectors of $\mathbf{v}_p$ and $\mathbf{v}_o$, respectively, a system of recurrence equations for animal i can be utilized (Hoeschele, 1993), or

$$\mathbf{g}_{o(i)} = \mathbf{b}_i^{*'} \mathbf{g}_{p(i)} + \boldsymbol{\varepsilon}_{(i)} \qquad [15]$$

with

$$\mathbf{b}_i^* = [\mathrm{Var}(\mathbf{g}_{p(i)})]^{-1}\mathrm{Cov}(\mathbf{g}_{o(i)}, \ \mathbf{g}_{p(i)}) \qquad [16]$$

where $\mathbf{b}_i^*$ and $\boldsymbol{\varepsilon}_{(i)}$ are vectors of corresponding partial regression coefficients and residual effects, respectively. Then the vector $\boldsymbol{\varepsilon}_{(i)}$ in equation [15] can be partitioned into the residuals of $\mathbf{a}_{o(i)}$ and $\mathbf{v}_{o(i)}^{*k}(k = 1 \text{ or } 2)$, or $\mathbf{m}^*$ and $\boldsymbol{\varepsilon}^{*k}$, which are uncorrelated, and the dispersion parameters for these vectors are given by the diagonal matrix $\mathbf{D}^*$ and the block diagonal matrix $\mathbf{G}_\varepsilon^*$, respectively. For animal i, the diagonal of $\mathbf{D}^*$ and the block of $\mathbf{G}_\varepsilon^*$ can be calculated by

$$d_{(i)}^* = \text{Var}(a_{o(i)}) - \mathbf{b}_i^{*'}\text{Var}(a_{p(i)})\mathbf{b}_i^* \qquad [17]$$

and

$$\mathbf{G}_{\varepsilon(i)}^* = \text{Var}\begin{pmatrix} \varepsilon_{(i)}^{*1} \\ \varepsilon_{(i)}^{*2} \end{pmatrix} = \text{Var}(\mathbf{v}_{o(i)}^*) - \mathbf{B}_i^*\text{Var}(\mathbf{v}_{p(i)}^*)\mathbf{B}_i^{*'} \qquad [18]$$

with the definition being provided later in the section *Computing dispersion parameters of residual effects*, where the matrix $\mathbf{B}_i^*$ relates $\mathbf{v}_{o(i)}^*$ to $\mathbf{v}_{p(i)}^*$ and has at most four non-zero elements of the conditional probabilities, as in tables I or II of Hoeschele (1993). Therefore the dispersion parameters of the residual effects in equation [7] must be replaced by

$$\mathbf{R}^* = \begin{bmatrix} \mathbf{I}_p\sigma_e^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* + \mathbf{P}_o\mathbf{G}_\varepsilon^*\mathbf{P}_o' + \mathbf{I}_o\sigma_e^2 \end{bmatrix} \qquad [19]$$

Consequently, the MMEs are given by

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{*-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{*-1}\mathbf{Z}_t & \mathbf{X}'\mathbf{R}^{*-1}\mathbf{L}^* \\ \mathbf{Z}_t'\mathbf{R}^{*-1}\mathbf{X} & \begin{bmatrix} \mathbf{Z}_t'\mathbf{R}^{*-1}\mathbf{Z}_t & \mathbf{Z}_t'\mathbf{R}^{*-1}\mathbf{L}^* \\ \mathbf{L}^{*'}\mathbf{R}^{*-1}\mathbf{Z}_t & \mathbf{L}^{*'}\mathbf{R}^{*-1}\mathbf{L}^* \end{bmatrix} + \mathbf{G}_p^{*-1} \\ \mathbf{L}^{*'}\mathbf{R}^{*-1}\mathbf{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}^o \\ \widehat{\mathbf{a}}_p \\ \widehat{\mathbf{v}}_p^* \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{*-1}\mathbf{y} \\ \mathbf{Z}_t'\mathbf{R}^{*-1}\mathbf{y} \\ \mathbf{L}^{*'}\mathbf{R}^{*-1}\mathbf{y} \end{bmatrix} \qquad [20]$$

where $\mathbf{G}_p^{*-1}$ is the inverse covariance matrix of the total additive genetic effects and the needed additive effects of MQTL alleles for the parent animals, which is computed according to the methodology as described by Hoeschele (1993), and $\mathbf{L}^* = -\mathbf{Z}_t\mathbf{P}_p^* + \mathbf{W}^*$. The matrices with asterisks represent the appropriate submatrices of the corresponding matrices in equation [14].

Backsolving for non-parent animals

The additive genetic effects and needed additive effects of MQTL alleles for non-parent animals can be computed by

$$\widehat{\mathbf{a}}_o = \mathbf{T}\widehat{\mathbf{a}_p} + \mathbf{L}_o^*\widehat{\mathbf{v}_p^*} + \widehat{\mathbf{m}^*} + \mathbf{P}_o^*\widehat{\boldsymbol{\varepsilon}^*} \qquad [21]$$

and

$$\widehat{\mathbf{v}_o} = \mathbf{B}^*\widehat{\mathbf{v}_p^*} + \widehat{\boldsymbol{\varepsilon}^*} \qquad [22]$$

where $\mathbf{L}_o^*$ is the appropriate submatrix of $\mathbf{L}^*$, and the vectors $\widehat{\mathbf{m}^*}$ and $\widehat{\boldsymbol{\varepsilon}^*}$ are given by

$$\begin{bmatrix} \widehat{\mathbf{m}^*} \\ \widehat{\boldsymbol{\varepsilon}^*} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_o + \mathbf{D}^{*-1}\sigma_e^2 & \mathbf{P}_o^* \\ \mathbf{P}_o^{*\prime} & \mathbf{P}_o^{*\prime}\mathbf{P}_o^* + \mathbf{G}_\varepsilon^{*-1}\sigma_e^2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}_o - \mathbf{X}_o\boldsymbol{\beta}^o - \mathbf{T}\widehat{\mathbf{a}_p} - \mathbf{L}_o^*\widehat{\mathbf{v}_p^*} \\ \mathbf{P}_o^{*\prime}(\mathbf{y}_o - \mathbf{X}_o\boldsymbol{\beta}^o - \mathbf{T}\widehat{\mathbf{a}_p} - \mathbf{L}_o^*\widehat{\mathbf{v}_p^*}) \end{bmatrix}$$

[23]

with equations [17] and [18].

Computing dispersion parameters of residual effects

First, as stated by Hoeschele (1993), for each marker used in the population, the needed additive effects of MQTL alleles are determined, and the list of the genotyped animals is created; this is referred to as the marker file. The following rules presented by Hoeschele (1993) can be applied to computing the matrices $\mathbf{D}^*$ and $\mathbf{G}_\varepsilon^*$.

For the matrix $\mathbf{D}^*$, if an animal i has one or both of its parents with the known marker data in the marker file, $\mathbf{D}^*$ in equation [19] equals $\mathbf{D}\sigma_u^2$ in equation [7]. For an animal i with the both parents in the pedigree file, if one or both of the parents are not retained in the marker file, the regression coefficients $\mathbf{b}_i^*$ in equation [17] are computed by equation [16] with equations [26] and [27] or with equations [28] and [29] in Hoeschele (1993), respectively, and then $d_{(i)}^*$ can be computed by equation [17] with $\mathrm{Var}(a_{o(i)}) = \sigma_a^2$, where $\mathbf{g}_{p(i)}$ and $a_{o(i)}$ are equivalent to $\mathbf{g}_{i,\mathrm{par}}$ and $\mathbf{a}_i$ in Hoeschele (1993), respectively. Also, if one of the parents (eg, s) is retained in the marker file, and the additive effect of the MQTL allele (eg, v_i^2) derived from another parent (eg, d) which is not retained in the marker file may be eliminated, then $d_{(i)}^*$ equals that given as equation [25] in Hoeschele (1993).

For the matrix $\mathbf{G}_\varepsilon^*$, if both parents of animal i are retained in the marker file, it can be calculated by equation [18]. With no inbreeding, then $\mathrm{Var}(\mathbf{v}_{o(i)}^*) = \mathbf{I}_{(2 \times 2)}\sigma_v^2$ and $\mathrm{Var}(\mathbf{v}_{p(i)}^*) = \mathbf{I}_{(4 \times 4)}\sigma_v^2$, where the subscripts represent the size of the identity matrices. With inbreeding, however, the matrix $\mathbf{G}_\varepsilon^*$ equals $\mathbf{G}_\varepsilon\sigma_e^2$ in equation [7], and $\mathrm{Var}(v_{s(i)}^1 \; v_{s(i)}^2 \; v_{d(i)}^1 \; v_{d(i)}^2 \; v_i^1 \; v_i^2)$ must be computed from a list of the additive effects of MQTL alleles for all animals in the pedigree file as described by Wang et al (1995). If both parents of animal i are known, and only one of them (eg, d) is retained in the marker file, we have

$$\mathbf{G}_{\varepsilon(i)}^* = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}\sigma_a^2 - \mathbf{B}_i^* \begin{bmatrix} t & tf_s & 0 \\ & t & 0 \\ \mathrm{Sym} & & (1 + F_d) \end{bmatrix}\mathbf{B}_i^{*\prime}\sigma_a^2$$

as given by Hoeschele (1993), where $t = \sigma_v^2/\sigma_a^2$, and f_s is the conditional inbreeding coefficient of sire for the MQTL, given the observed marker information, as described by Wang et al (1995). If the MQTL effects of both parents are eliminated, we have

$$\mathbf{G}_{\varepsilon(i)}^* = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}\sigma_a^2 - \mathbf{B}_i^* \begin{bmatrix} 1 + F_s & 0 \\ 0 & 1 + F_d \end{bmatrix}\mathbf{B}_i^{*\prime}\sigma_a^2$$

If one parent of an animal i is unknown, rows and columns pertaining to this parent in $\mathrm{Var}(\mathbf{v}^*_{p(i)})$ are deleted.

EXAMPLE

Six animals (animals 1–6) are considered for an illustration. Marker information on the animals is given in figure 1. Animal 1 has no record, and animals 2–6 have only one record each. The vector of records for animals 2–6 is [80 120 90 110 115]. Since parent animals 2 and 3 have no marker data, the additive effects of MQTL alleles for these animals can be eliminated. In addition, non-parent animal 5 is also eliminated, since the additive effect of an MQTL allele linked to M_4 was derived from its parent animals 3. Therefore, the needed additive effects of MQTL alleles are v_1^k and $v_4^k (k = 1$ or $2)$ for parent animals and v_5^1 and v_6^k for non-parent animals, where v_i^k represent the additive effects of MQTL alleles for animal i.
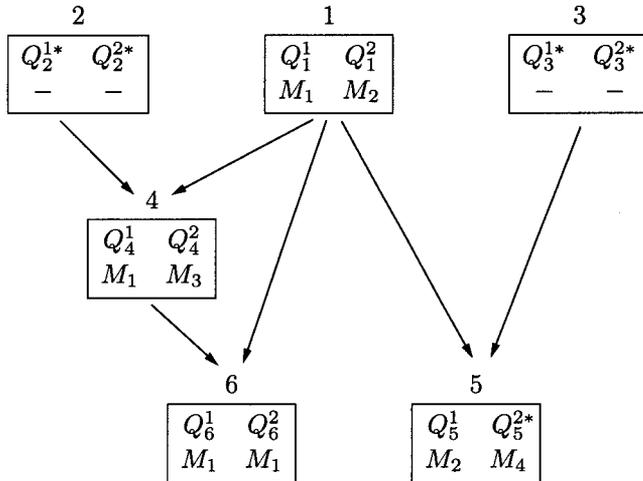

Fig 1. Example data for six animals with partial marker information. The symbols Q_i^k and M_j denote the MQTL and ML alleles, where the subscripts j represent the marker genotypes. The additive effects of MQTL alleles with asterisks can be eliminated in the current RAM approach.

In this paper, $r = 0.05$ is the recombination rate assumed between the ML and the MQTL. The assumed values for variance components are $\sigma_u^2 = 0.9$, $\sigma_v^2 = 0.05$, $\sigma_a^2 = 1$ and $\sigma_e^2 = 1.5$. The inverse covariance matrix of the total additive genetic effects and the needed MQTL effects for parent animals are given in table I, which are computed as described by Hoeschele (1993). The gametic relationship matrix for additive effects of MQTL alleles for all six animals is presented in table II.

The incidence matrix for the fixed effects is assumed to be

$$\mathbf{X}' = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Table I. The inverse covariance matrix of the total additive genetic effects and the needed additive effects of MQTL alleles for parent animals.

	Animal				MQTL effect		
a_1	a_2	a_3	a_4	v_1^1	v_1^2	v_4^1	v_4^2
1.652 471	0.512 8844	0	−1.082 72	−1.652 471	−1.652 471	1.082 7198	0.569 8354
	1.538 5382	0	−1.025 769	−0.512 884	−0.512 884	1.025 7687	−0.512 769
		1	0	0	0	0	0
			2.165 4396	1.082 7198	1.082 7198	−2.165 44	−1.139 671
				211.652 47	11.652 471	−201.0827	−0.569 835
	Sym				22.178 787	−11.609 04	−0.569 835
						212.691 76	1.139 6709
							21.652 44

Table II. The gametic relationship matrix of the additive effects of MQTL alleles for six animals.

v_1^1	v_1^2	v_2^1	v_2^2	v_3^1	v_3^2	v_4^1	v_4^2	v_5^1	v_5^2	v_6^1	v_6^2
1	0	0	0	0	0	0.95	0	0.05	0	0.926 25	0.926 25
	1	0	0	0	0	0.05	0	0.95	0	0.048 75	0.048 75
		1	0	0	0	0	0.5	0	0	0.0125	0.0125
			1	0	0	0	0.5	0	0	0.0125	0.0125
				1	0	0	0	0	0.5	0	0
					1	0	0	0	0.5	0	0
						1	0	0.095	0	0.9275	0.9275
							1	0	0	0.025	0.025
	Sym							1	0	0.092 625	0.092 625
									1	0	0
										1	0.859 75
											1

The matrices $\mathbf{Z_t}$ and $\mathbf{L}^*$ in equation [20] are written as

$$\mathbf{Z_t} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

and

$$\mathbf{L}^* = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.45 & 0.45 & 0 & 0 \\ 0.45 & -0.45 & 0.45 & -0.45 \end{bmatrix}$$

For equation [19], the matrix $\mathbf{R}^*$ is given as

$$\mathbf{R}^* = \mathrm{diag}\,[1.5 \quad 1.5 \quad 1.5 \quad 1.979\,75 \quad 1.9595]$$

where for the non-parent animals 5 and 6

$$\mathbf{D}^* = \mathrm{diag}\,[0.475 \quad 0.45]$$

and

$$\mathbf{G}_\varepsilon^* = \begin{bmatrix} 0.004\,75 & 0 & 0 \\ 0 & 0.005\,8813 & -0.001\,131 \\ 0 & -0.001\,131 & 0.005\,8813 \end{bmatrix}$$

Therefore, each effect for the parent animals can be obtained by solving the following MME (see next page) given as equation [20] with the matrices mentioned above. Moreover, using equations [21] and [22] with $\widehat{\mathbf{m}^*}$ and $\widehat{\varepsilon^*}$ given as equation [23], each effect of the non-parent animals is given. The $\widehat{\mathbf{m}^*}$ and $\widehat{\varepsilon^*}$ for this example are

$$\begin{bmatrix} \widehat{\mathbf{m}^*} \\ \widehat{\varepsilon^*} \end{bmatrix} = \begin{bmatrix} -1.291\,04 \\ 1.223\,09 \\ -0.012\,91 \\ 0.012\,91 \\ 0.012\,91 \end{bmatrix}$$

The solutions for fixed effects are 111.7257 and 97.9823, and those for the random effects are listed in table III, together with the estimates from the AM approach of Fernando and Grossman (1989).

The number of equations in the AMs of Fernando and Grossman (1989) and Hoeschele (1993) and in the RAM of Cantet and Smith (1991) are 20, 15 and 14, respectively, while that in the current RAM approach is ten. The solutions obtained by the current approach are equal to the corresponding ones in those methods.

Table III. The solutions for random effects.

Animal	Current RAM			AM [†]		
	a_i	v_i^1	v_i^2	u_i	v_i^1	v_i^2
1	−0.4533	0.1315	−0.1218	−0.4630	0.1315	−0.1218
2	−7.5383	–	–	−6.7294	−0.4044	−0.4044
3	7.9916	–	–	7.1925	0.3995	0.3995
4	−4.4003	0.1197	−0.4607	−4.0593	0.1197	−0.4607
5	2.3512	−0.1220	–	2.1416	−0.1220	0.3316
6	−0.8026	0.1177	0.1177	−1.0381	0.1177	0.1177

[†] The AM approach of Fernando and Grossman (1989).

Mixed-model equations for the example in figure 1.

$$
\begin{bmatrix}
1.01546 & 0 & 0.507\,72 & 0 & 0.252\,56 & 0.255\,17 & 0.002\,35 & -0.002\,35 & 0.229\,65 & -0.229\,65 \\
 & 2 & 0 & 0.666\,67 & 0.666\,67 & 0.666\,67 & 0 & 0 & 0 & 0 \\
 & & 1.906\,33 & 0.512\,88 & 0.126\,28 & -0.955\,14 & -1.651\,30 & -1.653\,65 & 1.197\,55 & 0.455\,01 \\
 & & & 2.205\,20 & 0 & -1.025\,77 & -0.512\,88 & -0.512\,88 & 1.025\,77 & -0.512\,77 \\
 & & & & 1.792\,95 & 0 & -0.113\,65 & 0.113\,65 & 0 & 0 \\
 & & & & & 2.959\,69 & 1.197\,55 & 0.967\,90 & -2.050\,61 & -1.254\,50 \\
 & & \text{Sym} & & & & 211.858\,1 & 11.446\,84 & -200.979\,4 & -0.673\,18 \\
 & & & & & & & 22.384\,42 & -11.712\,38 & -0.466\,49 \\
 & & & & & & & & 212.795\,1 & 1.036\,33 \\
 & & & & & & & & & 21.755\,78
\end{bmatrix}
\begin{bmatrix}
\beta_I^o \\ \beta_{II}^o \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\alpha}_4 \\ \hat{v}_1^1 \\ \hat{v}_1^2 \\ \hat{v}_4^1 \\ \hat{v}_4^2
\end{bmatrix}
=
\begin{bmatrix}
114.251\,01 \\ 193.333\,33 \\ 57.125\,51 \\ 53.333\,33 \\ 107.781\,29 \\ 89.344\,22 \\ 1.406\,64 \\ -1.406\,64 \\ 26.409\,80 \\ -26.409\,80
\end{bmatrix}
$$

DISCUSSION

The application of BLUP to MAS is expected to be useful for accelerating genetic progress through increasing accuracy of selection, reducing generation interval and increasing selection differential (eg, Soller, 1978; Soller and Beckmann, 1983; Smith and Simpson, 1986; Kashi et al, 1990; Meuwissen and van Arendonk, 1992; Ruane and Colleau, 1995). The RAM approach presented permits the best linear unbiased estimation (BLUE) of fixed effects and simultaneous BLUP of the total additive genetic merits for parent animals and the additive effects of the QTL alleles linked to the marker alleles only for animals that have known marker data or provide relationship ties among at least two descendants with known marker data. The current model is the RAM version of the AM derived by Hoeschele (1993) and is also equivalent to the AM of Fernando and Grossman (1989), which allows BLUE of fixed effects and simultaneous BLUP of the additive genetic effects due to QTL unlinked to the ML and the additive effects due to MQTL alleles.

Genetic evaluation of animals in the current population often requires information on their ancestors in the analysis. On the other hand, marker information will only be available on current animals, probably a limited number of elite animals even in the near future. Therefore, a substantial fraction of animals considered in the analysis may not be marker genotyped. Hoeschele's (1993) approach is attractive for these situations, since for the MQTL effects it requires the equations only for genotyped animals and ancestor animals connecting between any two animals with marker information provided. Thus, if many of the animals in the population do not have the marker data, the equations for the additive effects of MQTL alleles for such animals may not be needed with the current RAM approach. Moreover, since for the random effects, only the equations for parent animals are required, the number of equations may be considerably reduced in the current approach.

REFERENCES

Cantet RJC, Smith C (1991) Reduced animal model for marker assisted selection using best linear unbiased prediction. *Genet Sel Evol* 23, 221-233

Fernando RL, Grossman M (1989) Marker assisted selection using best linear unbiased prediction. *Genet Sel Evol* 21, 467-477

Goddard ME (1992) A mixed model for analyses of data on multiple genetic markers. *Theor Appl Genet* 83, 878-886

Grignola FE, Hoeschele I, Meyer K (1994) Empirical best linear unbiased prediction to map QTL. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, University of Guelph, Guelph, Canada, 21, 245-248

Henderson CR (1973) Sire evaluation and genetic trend. 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, USA, 10-41

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423-447

Henderson CR (1984) *Applications of Linear Models in Animal Breeding.* University of Guelph, Guelph, Ontario, Canada

Hoeschele I (1993) Elimination of quantitative trait loci equations in an animal model incorporating genetic marker data. *J Dairy Sci* 76, 1693-1713

Kashi Y, Hallerman E, Soller M (1990) Marker-assisted selection of candidate bulls for progeny testing programmes. *Anim Prod* 51, 63-74

Meuwissen THE, van Arendonk JAM (1992) Potential improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. *J Dairy Sci* 75, 1651-1659

Ruane J, Colleau JJ (1995) Marker-assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genet Res Camb* 66, 71-83

Saito S, Iwaisaki H (1996) The (co)variance structure of residual effects in a reduced animal model for marker-assisted selection. *Exp Anim* 45 (suppl), (abstr)

Saito S, Iwaisaki H (1997) A reduced animal model approach to predicting the total additive genetic merits for marker assisted selection. *Genet Sel Evol* 29 (in press)

Smith C, Simpson SP (1986) The use of genetic polymorphisms in livestock improvement. *J Anim Breed 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 JS (1983) Genetic polymorphism in varietal identification and genetic improvement. *Theor Appl Genet* 67, 25-33

van Arendonk JAM, Tier B, Kinghorn BP (1993) Simultaneous estimation of effects of unlinked markers and polygenes on a trait showing quantitative genetic variation. In: *Proceedings of the 17th International Congress on Genetics, Birmingham*, UK, 192

van Arendonk JAM, Tier B, Kinghorn BP (1994) Use of multiple genetic markers in prediction of breeding values. *Genetics* 137, 319-329

Wang T, Fernando RL, van der Beek S, Grossman M, van Arendonk JAM (1995) Covariance between relatives for a marked quantitative trait locus. *Genet Sel Evol* 27, 251-274

Weller JI, Fernando RL (1991) Strategies for the improvement of animal production using marker-assisted selection. In: *Gene Mapping: Strategies, Techniques and Applications* (LB Schook, HA Lewin, DG McLaren, eds), Marcel Dekker, New York, USA, 305-328