

# Efficiency of multistage marker-assisted selection in the improvement of multiple quantitative traits

CHONGQING XIE\* & SHIZHONG XU

*Department of Botany and Plant Science, University of California, Riverside, CA 92521–0124, U.S.A.*

The application of marker-assisted selection (MAS) to breeding programmes depends on its relative cost and the expected economic return compared to conventional phenotypic selection. The relative efficiency of MAS can be increased through a two-stage selection scheme or through marker-based, multiple-trait improvement. However, the effectiveness of these alternatives has not been quantified. In this study, we evaluate the efficiency of MAS relative to conventional phenotypic selection and marker-only selection in multistage selection for the improvement of multiple traits. We further incorporate the costs of obtaining measurements on phenotypic characters and marker loci into the objective function to evaluate the efficiency of MAS with respect to the gain per unit cost. Deterministic analyses indicate that excluding costs, multiple-trait MAS can be used to increase the aggregate breeding values in quantitative characters and is expected to be more effective than conventional selection or single-trait MAS. Two-stage MAS has a slightly reduced gain because of culling in the first stage. If the objective function is to maximize the gain per unit cost, multiple-trait MAS is inferior to phenotypic selection in most of the selection schemes investigated when the cost ratio ( $r$ ) of obtaining measurements on phenotypic characters to scoring marker loci is less than unity ( $r \leq 1.0$ ) and the heritability ( $h^2$ ) is greater than 0.3. The efficiency of MAS increases as  $r$  increases and  $h^2$  decreases. For MAS to be more effective, it is necessary to decrease further the cost associated with molecular marker assays.

**Keywords:** cost function, genetic gain, marker-assisted selection, multistage selection, selection index.

## Introduction

Rapid progress with molecular marker polymorphisms has made it possible for the first time to map the entire genome of virtually any species (Helentjaris *et al.*, 1986; Stuber, 1992). Because markers are unlimited in number and genomic distribution, this has renewed interest in the use of molecular markers to facilitate the identification of and selection for individual quantitative trait loci (QTLs) that control economically important traits. The steps in marker-assisted selection (MAS) consist of identifying associations between marker alleles and QTLs, or ideally, of estimating the contribution of marker loci to the genotypic value of the trait by the QTLs associated with the markers (MQTL effects), and combining these marker effects with phenotypic information to rank individuals through an index

and develop desired lines or populations (Lande & Thompson, 1990; Dudley, 1993; Gimelfarb & Lande, 1994).

MQTL effects can be identified and estimated through linkage disequilibrium (LD) created by crossing two inbred lines or divergent populations (Lande & Thompson, 1990), or through LD created within families in outbred species (Ruane & Colleau, 1995). Recent studies of MAS in breeding have tended to focus on the use of multiple regression of the phenotype on markers as a global method to identify markers linked to QTLs and to estimate marker effects (Lande & Thompson, 1990; Meuwissen & Van Arendonk, 1992; Zhang & Smith, 1992; Gimelfarb & Lande, 1994). Multiple regression has a computational advantage over the maximum likelihood method, while still producing very similar results (Lande & Thompson, 1990; Haley & Knott, 1992; Ruane & Colleau, 1995).

\*Correspondence. E-mail: cxie@evolution.ucr.edu

Lande & Thompson (1990; also Lande, 1992) proposed a theory of marker index selection (MIS) that maximizes the rate of genetic improvement under MAS by combining information on genetic marker polymorphism with data on phenotypic variation among individuals. Gimelfarb & Lande (1994) conducted simulation studies to test the theory and showed that the efficiency of MAS depends on several factors, including: the total number of markers in the genome; the number of markers contributing to the index; the population size; and the heritability of the character. Dudley (1993) discussed the issues on MAS and described selection using markers only. Knapp (1994) reviewed the difference between marker-only selection and MIS and described the method of estimating MQTL effects using ANOVA.

Numerous applications of MAS to breeding have been proposed in the literature. MAS can be used to assist in selecting parents, increasing the effectiveness of back-cross breeding, improving sex-limited traits, speeding the development of superior lines and populations by marker-based seedling assays, and increasing the efficiency of selection by eliminating expensive, slow or difficult trait assays (Tanksley *et al.*, 1981; Edwards *et al.*, 1987; Lande & Thompson, 1990; Stuber, 1992; Dudley, 1993; Knapp, 1994). However, MAS usually deals with one-stage single-trait improvement. Lande & Thompson (1990) implied that the efficiency of MAS can be increased through a two-stage selection of immatures (seedlings, embryos), first based on marker loci followed by conventional phenotypic selection, or through marker-based multiple-trait improvement (see also Dudley, 1993). However, the efficiencies of these alternatives relative to conventional selection procedures have not been quantified.

The objective of this study is not to detect MQTL associations or estimate MQTL effects, but to investigate the efficiency of multistage MAS as compared to conventional selection methods, under the assumption that MQTL effects and index parameters are known. We also incorporate the costs associated with measuring phenotypic characters and scoring marker loci into the objective function to maximize the gain per unit cost.

## Improvement of a single trait

### *One-stage selection*

Lande & Thompson (1990; also Lande, 1992) described the theory of marker index selection. Let  $\mathbf{y}$  be a  $2 \times 1$  vector that contains a net molecular score

( $m$ ), which is the sum of the additive effects on the character for any individual associated with these markers and individual phenotypic value ( $x$ ), i.e.  $\mathbf{y} = [m \ x]^T$ . The two components have relative economic weights of  $\mathbf{w} = [0 \ 1]^T$ . Let  $\mathbf{P}$  and  $\mathbf{G}$  denote the phenotypic and genotypic variance-covariance matrices of vector  $\mathbf{y}$ , respectively. The MAS index proposed by Lande & Thompson (1990) is:

$$I = b_m m + b_x x,$$

where  $\mathbf{b}$  is the index coefficients derived from classical selection index theory; i.e.  $\mathbf{b} = [b_m \ b_x]^T = \mathbf{P}^{-1} \mathbf{G} \mathbf{w}$ . This index can be used for either individual selection or selection of lines (Lande & Thompson, 1990; Knapp, 1994).

### *Two-stage selection*

Marker selection at an early stage followed by conventional phenotypic selection of surviving adults has received considerable attention (Tanksley *et al.*, 1981; Soller & Beckmann, 1983; Lande & Thompson, 1990). This sort of sequential selection will reduce the total genetic gain as compared to single-stage selection because some outstanding individuals may be culled at the marker selection stage (Xu & Muir, 1991, 1992). Gimelfarb & Lande (1994) stated that selection is more effective if markers contributing to the index are re-evaluated each generation. When the second-stage selection is based on an index including the molecular score, rather than the adult phenotype value alone, the genetic gain is hard to predict (Lande & Thompson, 1990). The problem is that the optimum culling strategy is difficult to find because numerical multiple integration is required (Xu & Muir, 1992).

In this study, Xu & Muir's (1991, 1992) algebra is followed to derive the formulae for two-stage index selection. The molecular score is selected at the first stage, whereas the adult phenotype together with the molecular score are selected at the second stage. The method for constructing indices is to find  $\mathbf{b}$  so that the correlation between indices ( $z_i$ ) and the aggregate breeding value,  $H = \mathbf{w}^T \mathbf{g}$ , is maximum under the constraint of  $\text{Cov}(z_i, z_j) = 0$ , for  $i \neq j$ . The constraint ensures the existence of an exact solution for truncation points without resorting to numerical multiple integration (Xu & Muir, 1991, 1992). As pointed out by those authors, the genetic gain will be slightly reduced using this index because the restriction of orthogonality among  $\mathbf{z} = [z_1 \dots z_n]^T$  produces an effect similar to that of a restricted selection index. However, Xu & Muir (1991) also

showed that, under certain conditions, the efficiency of transformed culling may greatly exceed that of conventional independent culling because the former incorporates the information from previous stages.

Suppose now that  $\mathbf{y} = [m \ x]^T$  is to be selected in two stages, and that the restriction is started at the second stage. Let  $\mathbf{z} = [z_1 \ z_2]^T$  be a  $2 \times 1$  vector of the updated selection indices defined by Xu & Muir (1992). At the first stage, selection is based on the marker score ( $m$ ) only, and thus  $z_1 = b_{11}m$ . At the second stage, selection is based on the character phenotype ( $x$ ) and the marker score. The index for the second stage has the form of  $z_2 = b_{12}m + b_{22}x$ . In matrix notation, the two-stage index coefficients have the form of:

$$\mathbf{B} = \begin{pmatrix} b_{11} & 0 \\ b_{12} & b_{22} \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{(\sigma_g^2 - \sigma_m^2)}{(\sigma_p^2 - \sigma_m^2)} & \frac{(\sigma_g^2 - \sigma_m^2)}{(\sigma_p^2 - \sigma_m^2)} \end{bmatrix}, \quad (1)$$

where  $\sigma_m^2$  is the additive genetic variance associated with the marker loci,  $\sigma_g^2$  and  $\sigma_p^2$  are the additive genetic and phenotypic variances of the character.

Matrix  $\mathbf{B}$  is the transformation matrix from  $\mathbf{y}$  to  $\mathbf{z}$ , i.e.  $\mathbf{z} = \mathbf{B}^T \mathbf{y}$ . The responses to selection are:

$$\Delta \mathbf{G} = \begin{pmatrix} \Delta G_m \\ \Delta G_x \end{pmatrix} = \begin{pmatrix} \sigma_m \Delta z_1 \\ \sigma_m \Delta z_1 + (\sigma_p^2 - \sigma_m^2)^{-1/2} (\sigma_g^2 - \sigma_m^2) \Delta z_2 \end{pmatrix}, \quad (2)$$

where  $\Delta z_1$  and  $\Delta z_2$  are standardized selection intensities for stages one and two, respectively.  $\Delta G_m = \Delta z_1 \sigma_m$  is the correlated response in the average adult phenotype to marker selection on the immatures.  $\Delta G_x$  has  $\Delta G_m$  as one component and is the response relating to selection on the character phenotype. The corresponding gain in the aggregate breeding value ( $\Delta H_{TM}$ ) is:

$$\Delta H_{TM} = \mathbf{w}^T \Delta \mathbf{G} = (0 \ 1)^T \begin{pmatrix} \Delta G_m \\ \Delta G_x \end{pmatrix} = \Delta G_x. \quad (3)$$

The efficiency of this two-stage MIS relative to conventional phenotypic selection on the adult phenotype is:

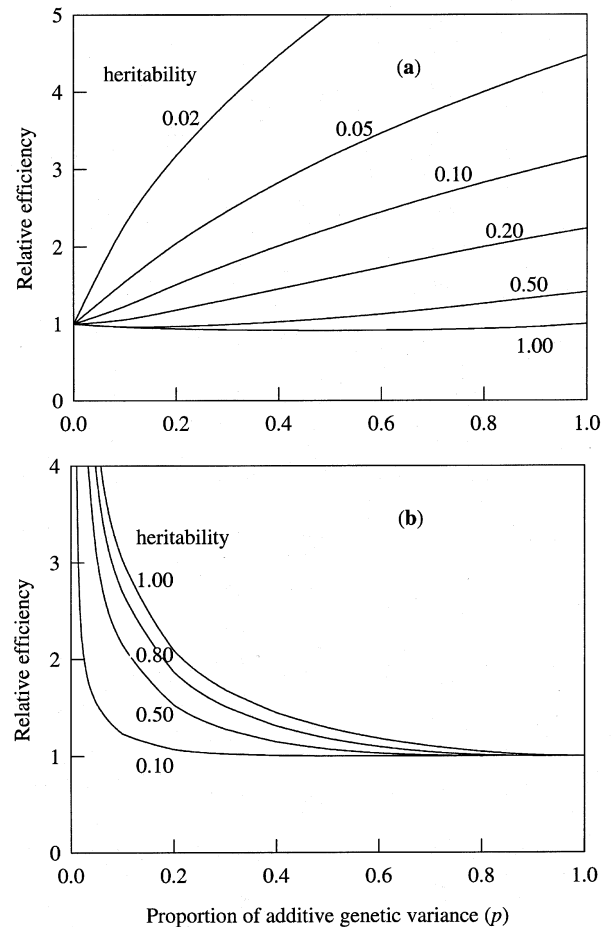
$$\Delta H_{TM} / \Delta H_P = \frac{1}{\Delta z_p} \left( \Delta z_1 \sqrt{\frac{p}{h^2}} + \Delta z_2 \sqrt{\frac{(1-p)^2}{1-ph^2}} \right), \quad (4)$$

where  $\Delta H_P = \Delta z_p h^2 \sigma_p$  is the aggregate genetic gain from phenotypic selection,  $h^2$  is the heritability of the character,  $p$  is the proportion of the additive genetic variance in the character that is associated

with the marker loci (Neimann-Sorensen & Robertson, 1961; Lande & Thompson, 1990) and  $\Delta z_i$  is the selection intensity of the  $i$ th stage.

The efficiency of two-stage MIS relative to conventional phenotypic selection with the same final proportion of 6.0% selected is depicted in Fig. 1(a) as a function of  $p$  for various values of  $h^2$ . When  $h^2 = 1.0$ , the two-stage MIS has no advantage over the phenotypic selection in terms of aggregate breeding values. The relative efficiency of MIS can be very large for a character with low heritability if a large proportion of the additive genetic variance is associated with the markers. The relative efficiency increases as  $p$  increases and  $h^2$  decreases.

The predicted aggregate genetic gain from marker-based selection on  $m$  is  $\Delta H_M = \Delta z_m \sigma_m$ . The efficiency with respect to the aggregate genetic gain



**Fig. 1** Efficiency of two-stage marker index selection (MIS) in the improvement of a single character, relative to conventional phenotypic selection (a) and to marker-only selection (b). The efficiency is plotted as a function of  $p$ , the proportion of additive variance associated with the marker loci, and for various  $h^2$  values.

of two-stage MIS relative to marker-only selection can be expressed as:

$$\Delta H_{TM}/\Delta H_M = \frac{1}{\Delta z_m} \left( \Delta z_1 + \Delta z_2 \frac{(1-p)h}{\sqrt{p(1-ph^2)}} \right). \quad (5)$$

The marker-only selection is more effective than phenotypic selection when  $\sigma_m^2/\sigma_g^2 > h^2$  (Smith, 1967; Dudley, 1993), but it is never superior to MIS (Knapp, 1994). Results in this study indicate that two-stage MIS with the same final proportion of 6.0% selected is more efficient than marker-only selection (Fig. 1b). The difference in efficiency decreases as  $p$  increases and  $h^2$  decreases. When  $p = 1$ , the two methods are equivalent. Knapp (1994) suggested that MIS should always be used when index parameters can be estimated. Otherwise the marker-only selection must be used when phenotypic index coefficients cannot be estimated, such as in unreplicated progenies. When  $p > 0.5$ , the marker-only selection is nearly as efficient as MIS selection. Thus, marker-only selection can be used to eliminate expensive, slow or difficult phenotypic trait assays particularly when  $h^2 < 0.5$  and  $p > 0.5$  (Fig. 1b).

#### Optimization with respect to economic gain per unit cost

The total genetic improvement on the average adult phenotype in a two-stage selection scheme is somewhat reduced as compared to that of single-stage selection. However, the advantage of two-stage selection is justified by cost savings associated with measuring traits because not all individuals need to be recorded for all traits (Cunningham, 1975; Xu & Muir, 1991, 1992). The procedure may be modified to obtain a maximum  $\Delta H$  per unit cost (Namkoong, 1970; Xu & Muir, 1992). Namkoong (1970) proposed a linear cost function for all individuals in a two-stage selection procedure. The linear cost function per individual ( $c$ ) for a two-stage selection can be expressed as:

$$c = c_1 + c_2 q_1,$$

where  $c_1$  and  $c_2$  are the costs of obtaining measurements on the marker loci and on the character phenotype, respectively, and  $q_1$  is the proportion selected at the first stage. The quantity to be maximized is:

$$Q = \Delta H/c$$

subject to the constraint  $q = q_1 q_2$ , where  $q$  is a predetermined total proportion selected. Because numerical integration is not required, the solution

can be easily obtained using the Newton–Raphson iterative equation system (Xu & Muir, 1992).

The relative efficiency with respect to the gain per unit cost of two-stage MIS to the conventional phenotypic selection is:

$$\frac{r}{\Delta z_p(1+rq_1)} \left( \Delta z_1 \sqrt{\frac{p}{h^2}} + \Delta z_2 \sqrt{\frac{(1-p)^2}{1-ph^2}} \right), \quad (6)$$

where  $r = c_2/c_1$ , the cost ratio. Under a predetermined proportion selected ( $q$ ),  $Q$  can be maximized by optimum allocation of  $\Delta z_1$  and  $\Delta z_2$ . The efficiency of two-stage MIS to phenotype selection is related to  $h^2$ ,  $p$ ,  $r$  and  $q_i$  (or  $\Delta z_i$ ).

Under the same final proportion selected, the efficiency of two-stage MIS in terms of gain per unit cost relative to conventional phenotypic selection is plotted in Fig. 2 as a function of  $p$  and  $r$  for  $h^2 = 0.1$  and  $h^2 = 0.3$ . Under the assumption that a molecular marker assay is more expensive than obtaining measurements on the character phenotype (i.e.  $r \leq 1$ ), the two-stage MIS for the improvement of a single trait is less efficient than conventional phenotypic selection when  $h^2 = 0.3$ . The efficiency increases as  $h^2$  decreases and  $r$  and  $p$  increase. It will be shown later that single-stage MIS with respect to the gain per unit cost is inferior to phenotypic selection because the former is never superior to two-stage MIS.

The efficiency of two-stage MIS in terms of the gain per unit cost relative to one-stage MIS is:

$$\frac{1+r}{\Delta z_p(1+rq_1)} \left( \frac{\Delta z_1 \sqrt{p-p^2h^2} + \Delta z_2 h(1-p)}{\sqrt{p-2ph^2+h^2}} \right). \quad (7)$$

This relative efficiency is plotted in Fig. 3 as a function of  $p$  and  $r$  for  $h^2 = 0.1$  and  $h^2 = 0.3$  under the same final proportion of 6.0% selected. The two-stage MIS is always superior to one-stage MIS unless  $r = 0$ . The efficiency increases as  $p$  and  $r$  increase but is weakly affected by  $h^2$ .

## Improvement of multiple traits

### One-stage selection

MAS in most cases deals with the improvement of a single trait. When marker effects are significant for more than one trait, MAS on multiple traits is expected to be more efficient in a multivariate context than in a univariate analysis of total economic value alone (Lande & Thompson, 1990; Dudley, 1993; Jiang & Zeng, 1995). This is because marker information can be simultaneously used to provide the MQTL effects for multiple traits. Extension of single-trait to multiple-trait MIS is straight-

forward. Lande & Thompson (1990) gave the general formula for the multiple-trait MAS index as  $I = \mathbf{b}_m^T \mathbf{m} + \mathbf{b}_x^T \mathbf{x}$  where:

$$\begin{pmatrix} \mathbf{b}_m \\ \mathbf{b}_y \end{pmatrix} = \begin{pmatrix} [\mathbf{I} - (\mathbf{P} - \mathbf{M})^{-1}(\mathbf{G} - \mathbf{M})]\mathbf{w} \\ (\mathbf{P} - \mathbf{M})^{-1}(\mathbf{G} - \mathbf{M})\mathbf{w} \end{pmatrix}.$$

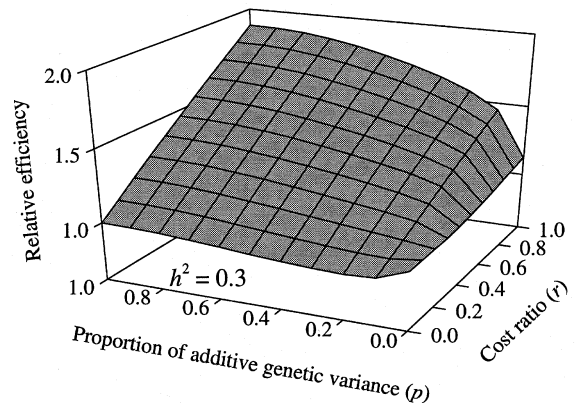
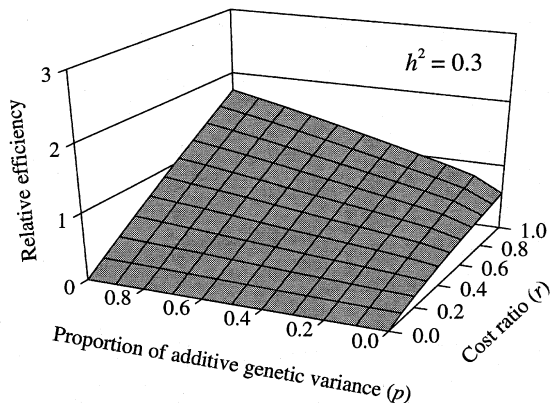
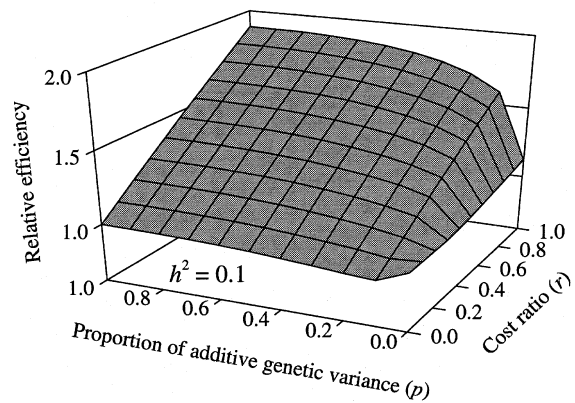
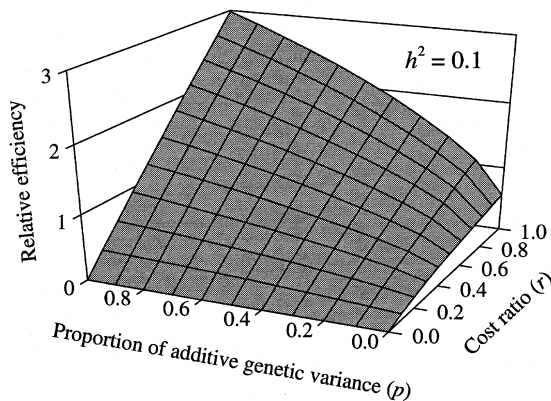
Note that  $\mathbf{x}$  is a vector of quantitative traits with phenotypic and additive genetic variance-covariance matrices  $\mathbf{P}$  and  $\mathbf{G}$ . Vector  $\mathbf{m}$  corresponds to the molecular scores obtained by summing the vectors of effects on the traits produced by associated molecular marker loci.  $\mathbf{M}$  is the covariance matrix between the breeding values and MQTL effects. The vector of relative economic weights for quantitative traits is  $\mathbf{w}$ , and that for the molecular markers is  $\mathbf{0}$ . Lande & Thompson (1990) showed that the relative

weights on the quantitative characters differ from those under purely phenotypic selection, and that the relative weights on the molecular scores are not proportional to simple economic values of the corresponding characters.

The relative efficiency of multiple-trait MIS over conventional index selection without marker information is:

$$\frac{\Delta z_I \{ \mathbf{w}^T [ (\mathbf{P} - \mathbf{M})^{-1} \mathbf{M} (\mathbf{P} - \mathbf{G}) + (\mathbf{P} - \mathbf{M})^{-1} \mathbf{G} (\mathbf{G} - \mathbf{M}) ] \mathbf{w} \}^{1/2}}{\Delta z_p (\mathbf{w}^T \mathbf{G} \mathbf{P}^{-1} \mathbf{G} \mathbf{w})^{1/2}}, \tag{8}$$

where  $\Delta z_I$  and  $\Delta z_p$  are the standardized selection intensity for MIS and phenotypic index selection, respectively, the numerator is the gain from MIS,



**Fig. 2** Efficiency of two-stage marker index selection (MIS) with respect to the gain per unit cost ( $\Delta H/c$ ) in the improvement of a single character, relative to conventional phenotypic selection. The efficiency is plotted as a function of  $p$ , the proportion of additive variance associated with the marker loci, and  $r$ , the ratio of the costs for measuring the phenotype and marker loci, and for heritability values of  $h^2 = 0.1$  (above) and  $h^2 = 0.3$  (below).

**Fig. 3** Efficiency of two-stage marker index selection (MIS) with respect to the gain per unit cost ( $\Delta H/c$ ) in the improvement of a single character, relative to one-stage MIS. The efficiency is plotted as a function of  $p$ , the proportion of additive variance associated with the marker loci, and  $r$ , the ratio of the costs for measuring phenotype and marker loci, and for heritability values of  $h^2 = 0.1$  (above) and  $h^2 = 0.3$  (below).

and the denominator is the gain from phenotypic index selection.

#### Two-stage selection

Multiple-trait MIS can be readily extended to the case of two-stage MIS, with the MQTL effects selected at the first stage and the phenotypic characters, together with the MQTL effects, selected at the second stage. Under the constraint that the indices at different stages are independent (Xu & Muir, 1991, 1992), the vector of index coefficients in a two-stage selection that maximizes the rate of genetic improvements is given by:

$$\mathbf{B} = \begin{pmatrix} \mathbf{w}^T & \mathbf{0}^T \\ [-\mathbf{(P-M)}^{-1}(\mathbf{G-M})\mathbf{w}]^T & [\mathbf{(P-M)}^{-1}(\mathbf{G-M})\mathbf{w}]^T \end{pmatrix} \quad (9)$$

(see Appendix A). The response to selection on this index is:

$$\Delta\mathbf{G} = \begin{pmatrix} \Delta\mathbf{G}_1 \\ \Delta\mathbf{G}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{M}^{1/2}\Delta z_1 \\ \mathbf{M}^{1/2}\Delta z_1 + \mathbf{(P-M)}^{-1/2}(\mathbf{G-M})\Delta z_2 \end{pmatrix}, \quad (10)$$

where  $\Delta z_1$  and  $\Delta z_2$  are the selection intensities for stages one and two.

The relative efficiency of two-stage MIS with respect to the gain per unit cost compared to phenotypic index selection is:

$$\frac{r}{(1+rq_1)} \left( \frac{\Delta H_{TM}}{\Delta H_P} \right), \quad (11)$$

where  $r$  is the ratio of the cost of obtaining measurements on all phenotypic characters to the cost of scoring marker loci,  $\Delta H_{TM} = \mathbf{w}^T[\mathbf{M}^{1/2}\Delta z_1 + \mathbf{(P-M)}^{-1/2}(\mathbf{G-M})\Delta z_2]$  is the aggregate genetic gain from two-stage MIS, and  $\Delta H_P = (\mathbf{w}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}\mathbf{w})^{1/2}\Delta z_p$  is the gain from phenotypic index selection.

The selection indices developed in this way are readily extended to more than two-stage selection (Appendix A). Because numerical integration is not involved in the calculation, the procedure can be used for any number of traits and for any reasonable number of stages without concerns about the amount of computation time required.

#### Numerical example

Two examples are given to illustrate multistage MIS for the improvement of multiple traits as compared with conventional index selection, marker-only selection and one-stage MIS with the same final proportion selected. Comparisons are made on the relative

efficiencies of aggregate genetic gains ( $\Delta H$ ) or the gain per unit cost ( $Q = \Delta H/c$ ) for the different selection procedures. Parameters investigated are heritability ( $h^2$ ), proportion of additive genetic variance associated with marker loci ( $p$ ), genotypic ( $r_g$ ) and phenotypic ( $r_p$ ) correlations, the cost ratio ( $r$ ) of measuring phenotypic traits to marker loci, and economic weights ( $w$ ). Because of the large number of possible combinations with four-trait selection, comparisons are limited to the cases of equal  $h^2$ ,  $r_g$  or  $r_p$ ,  $p$ , and  $r$  with several levels for each parameter and  $w = 1$  for each factor (Appendix B).

In two-stage MIS, MQTL effects are selected at the first stage, followed by MQTL effects plus phenotypic traits selected at the second stage. The results show that the relative efficiency with respect to the aggregate genetic gain of one- ( $R_{OM/P}$ ) and two-stage ( $R_{TM/P}$ ) MIS to conventional index selection for the improvement of two traits in all cases tested ranges from 100 to 278% and from 95 to 268%, respectively (Table 1). In general, two-stage MIS has relatively smaller aggregate genetic gains than one-stage MIS because of early culling. In all the cases, MIS is notably superior to marker-only selection.

The efficiency with respect to the gain per unit cost of two-stage MIS to conventional index selection ( $Q_{TM/P}$ ) for the improvement of two traits ranges from 52 to 301% (Table 2). The inferiority of MIS to conventional index selection occurs in the cases where  $r \leq 0.5$  (for two traits together,  $r \leq 1.5$ ). When  $r = 2$  (two traits together), MIS is superior to the phenotypic index selection. In addition, two-stage MIS in all the cases tested is notably superior to both one-stage MIS ( $Q_{TM/OM}$ ) and marker-only selection ( $Q_{TM/M}$ ).

In four-stage MIS for the genetic improvement of four traits, MQTL effects are selected at the first stage, followed by MQTL effects plus two, three and four phenotypic traits selected at the second, the third and the fourth stage, respectively. Results show that the efficiency of MIS to conventional index selection is even larger (Table 3). The four-stage MIS is superior to phenotypic index selection when  $r \geq 0.8$  (four traits together). The relative efficiency of MIS to the conventional index selection ranges from 54 to 506%. In these cases, the reduced gain in the four-stage MIS is offset by the cost savings arising from early culling.

#### Discussion

With the advent of new molecular technology, large numbers of marker loci can be determined in virtu-

**Table 1** Efficiency of single ( $\Delta H_{OM}$ ) or two-stage ( $\Delta H_{TM}$ ) marker index selection with respect to genetic gain for the improvement of two traits relative to phenotypic ( $R_{OM/P}$  or  $R_{TM/P}$ ) and to marker-only ( $R_{OM/M}$  or  $R_{TM/M}$ ) selection

No.	$r_p$	$r_g$	Trait 1			Trait 2			One-stage selection			Two-stage selection		
			$h^2$	$p$	$w$	$h^2$	$p$	$w$	$R_{OM/P}$	$R_{OM/M}$	$\Delta H_{OM}$	$R_{TM/P}$	$R_{TM/M}$	$\Delta H_{TM}$
1	0.2	0.2	0.6	0.8	1.0	0.3	0.2	1.0	1.121	1.183	1.625	1.036	1.093	1.502
2	0.2	0.2	0.3	0.2	0.2	0.6	0.8	1.0	1.100	1.183	1.348	1.017	1.093	1.247
3	0.2	0.8	0.1	0.2	1.0	0.6	0.8	1.0	1.098	1.160	1.762	1.019	1.076	1.634
4	0.2	0.8	0.6	0.8	1.0	0.1	0.2	0.2	1.095	1.164	1.405	1.016	1.079	1.302
5	0.8	0.2	0.6	0.2	1.0	0.1	0.8	1.0	1.026	3.013	1.823	0.980	2.879	1.742
6	0.8	0.2	0.1	0.8	1.0	0.6	0.2	0.2	1.141	1.706	0.394	1.049	1.568	0.363
7	0.2	0.2	0.3	0.2	1.0	0.3	0.2	1.0	1.051	2.304	0.969	0.985	2.161	0.909
8	0.2	0.2	0.3	0.4	1.0	0.3	0.4	1.0	1.105	1.681	1.019	1.015	1.544	0.936
9	0.2	0.8	0.1	0.2	0.2	0.1	0.2	1.0	1.057	2.651	0.294	1.001	2.511	0.278
10	0.2	0.8	0.6	0.2	0.2	0.6	0.2	1.0	1.003	2.516	1.673	0.947	2.375	1.579
11	0.8	0.2	0.1	0.8	0.2	0.1	0.8	1.0	1.216	1.559	0.305	1.114	1.428	0.280
12	0.8	0.2	0.6	0.8	0.2	0.6	0.8	1.0	2.784	3.568	4.195	2.687	3.445	4.050
13	0.2	0.2	0.1	0.2	1.0	0.1	0.2	1.0	1.081	2.371	0.333	1.016	2.228	0.313
14	0.2	0.2	0.6	0.2	1.0	0.6	0.2	1.0	1.018	2.232	1.878	0.953	2.089	1.758

**Table 2** Efficiency of two-stage marker index selection in terms of the gain per unit cost ( $Q_{TM}$ ) for the improvement of two traits relative to phenotypic ( $Q_{TM/P}$ ), marker-only ( $Q_{TM/M}$ ), and one-stage marker index ( $Q_{TM/OM}$ ) selection

No.	$r_p$	$r_g$	Trait 1				Trait 2				Relative efficiency			
			$h^2$	$p$	$r$	$w$	$h^2$	$p$	$r$	$w$	$Q_{TM/P}$	$Q_{TM/M}$	$Q_{TM/OM}$	$Q_{TM}$
1	0.2	0.2	0.6	0.8	1.0	1.0	0.3	0.2	1.0	1.0	1.751	2.771	2.343	1.269
2	0.2	0.2	0.3	0.2	0.5	0.2	0.6	0.8	1.0	1.0	1.334	2.390	2.021	1.090
3	0.8	0.5	0.6	0.2	0.5	0.2	0.3	0.8	1.0	1.0	1.284	2.614	1.872	0.556
4	0.8	0.5	0.1	0.4	0.2	1.0	0.3	0.2	0.5	1.0	0.528	3.129	1.235	0.602
5	0.2	0.8	0.1	0.2	1.0	1.0	0.6	0.8	1.0	1.0	1.735	2.751	2.371	1.392
6	0.2	0.8	0.6	0.8	1.0	1.0	0.1	0.2	0.5	0.2	1.341	2.373	2.039	1.146
7	0.8	0.2	0.6	0.2	1.0	1.0	0.1	0.8	1.0	1.0	1.150	5.068	1.682	1.022
8	0.8	0.2	0.1	0.8	1.0	1.0	0.6	0.2	0.5	0.2	1.193	2.972	1.742	0.275
9	0.2	0.2	0.3	0.2	0.5	1.0	0.3	0.2	0.5	1.0	0.729	3.199	1.389	0.673
10	0.2	0.2	0.3	0.4	0.5	1.0	0.3	0.4	0.5	1.0	0.824	2.506	1.492	0.760
11	0.8	0.8	0.3	0.4	1.0	0.2	0.3	0.4	1.0	1.0	1.534	3.204	2.046	0.533
12	0.8	0.8	0.3	0.8	1.0	0.2	0.3	0.8	1.0	1.0	2.062	2.751	2.371	0.716
13	0.2	0.8	0.1	0.2	0.5	0.2	0.1	0.2	0.5	1.0	0.715	3.587	1.353	0.199
14	0.2	0.8	0.6	0.2	0.5	0.2	0.6	0.2	0.5	1.0	0.685	3.437	1.366	1.142
15	0.8	0.2	0.1	0.8	1.0	0.2	0.1	0.8	1.0	1.0	1.662	3.195	2.050	0.209
16	0.8	0.2	0.6	0.8	1.0	0.2	0.6	0.8	1.0	1.0	3.015	5.798	1.624	2.272
17	0.2	0.2	0.1	0.2	0.5	1.0	0.1	0.2	1.0	1.0	1.032	3.771	1.591	0.212
18	0.2	0.2	0.6	0.2	0.5	1.0	0.6	0.2	1.0	1.0	0.986	3.603	1.615	1.213

ally any species. Researchers can now routinely score large numbers of such polymorphic loci on many individuals in a population and determine the architecture of quantitative traits with unprecedented precision (Stuber, 1992; Dudley, 1993).

However, few breeders appear to adopt MAS in applied breeding programmes. The main impediment is cost (Lande, 1992; Strauss *et al.*, 1992; Dudley, 1993). Despite numerous improvements made in the last decade, scoring molecular markers

**Table 3** Efficiency of four-stage marker index selection in terms of the gain per unit cost ( $Q_{MI}$ ) for the improvement of four traits relative to phenotypic index ( $Q_{MI/P}$ ), marker-only ( $Q_{MI/M}$ ), and one-stage marker index ( $Q_{MI/OM}$ ) selection

Case	$r_p, r_g$	$h^2$	$p$	$r$	Relative efficiency			
					$Q_{MI/P}$	$Q_{MI/M}$	$Q_{MI/OM}$	$Q_{MI}$
1	0.2	0.1	0.8	0.5	2.070	2.853	2.212	0.520
2	0.2	0.3	0.2	0.2	0.592	2.816	1.264	1.116
3	0.2	0.3	0.8	0.5	1.996	2.751	2.331	1.503
4	0.8	0.1	0.8	0.5	2.794	2.742	2.446	1.023
5	0.8	0.3	0.8	0.2	1.167	1.718	1.678	3.205
6	0.8	0.3	0.2	1.0	2.429	5.582	2.833	1.334
7	-0.2	0.1	0.8	0.5	1.405	3.268	1.822	0.176
8	-0.2	0.3	0.2	0.2	0.544	2.919	1.174	0.513
9	-0.2	0.3	0.8	0.5	1.365	3.175	1.860	0.514
10	-0.2	0.3	0.8	1.0	2.335	4.526	2.652	0.440
11	0.2	0.6	0.2	0.5	1.236	3.919	1.820	1.863
12	0.2	0.3	0.8	1.0	3.570	4.100	3.475	1.345
13	0.2	0.6	0.8	0.2	0.833	1.722	1.625	3.137
14	0.8	0.6	0.2	0.2	0.640	2.647	1.403	3.515
15	0.8	0.6	0.8	1.0	5.061	4.139	3.456	5.558

is more expensive than obtaining measurements on phenotypic characters (Lande, 1992). The application of MAS to applied breeding programmes depends on the relative cost and the expected return compared to conventional breeding methods. In this study, we incorporated the costs associated with collection of data into the objective function to maximize the gain per unit cost. The results of MAS with respect to the gain per unit cost, as compared with conventional phenotypic selection, are discouraging unless new progress is made in further reducing the costs associated with molecular marker assays.

Throughout this paper, a number of assumptions have been made. First, a large sample size is assumed in the calculation of index parameters. The significance of increasing sample size on MAS is twofold: it increases the proportion of the additive genetic variance in a character likely to be detected through MQTL associations; and it reduces sampling error (Lande & Thompson, 1990). Secondly, MQTL effects and index parameters are assumed to be known. We are focusing on evaluating the relative efficiency of multistage MIS for the improvement of multiple traits with respect to aggregate genetic gain ( $\Delta H$ ) or the gain per unit cost ( $\Delta H/c$ ). Thirdly, a number of factors that may favour MAS have not been included in the comparison. These

factors include pollen control coefficients, selection cycle length and relative selection intensity (Fehr, 1987; Dudley, 1993; Edwards & Page, 1994).

An optimum multistage index selection can be performed even if the index parameters will be estimated simultaneously. Consider a two-stage MAS for a single trait; the markers are selected at the first stage, whereas the adult phenotype, together with the markers, are selected at the second stage. Essentially, the selection criterion at the first stage is based on the probability of allelic transmission at QTLs plus estimated breeding values of the polygenes. What the markers can do is to help us guess better which QTL allele has been transmitted to the offspring. The estimated breeding values of individuals surviving the initial cullings will be updated when the phenotypic records are available. Fernando & Grossman's (1989) mixed-model equations can be directly applied because the preliminary cullings can be incorporated into best linear unbiased prediction under the selection model (Xie & Xu, 1996). By measuring both the genetic gain and the cost consumed, an optimum selection index can be constructed.

Marker loci are expected to be significantly associated with many phenotypic characters (Edwards *et al.*, 1987; Stuber *et al.*, 1987; Stuber, 1992). In such a case, simultaneous improvement of several traits using a marker-based selection index can be achieved (Dudley, 1993; Jiang & Zeng, 1995). Alternatively, an index value, based on phenotypic characters and MQTL effects, is desirable (Lande & Thompson, 1990). For simultaneous improvement of several traits, deterministic analyses in this study indicate that multiple-trait MIS can be used to increase substantially the aggregate breeding value in quantitative characters. The potential efficiency of multiple-trait MIS relative to conventional index selection in terms of aggregate breeding values depends on heritability ( $h^2$ ), proportion of marker-associated additive genetic variance ( $p$ ), genotypic ( $r_p$ ) and phenotypic ( $r_g$ ) correlations, and economic weights ( $w$ ). The multiple-trait MIS is expected to be more effective than conventional phenotypic selection and marker-only selection. Unfortunately, it is impossible to estimate marker and phenotypic index parameters in unreplicated progeny tests. In these cases, marker-only selection can be used because it is nearly as efficient as MIS when  $p > 0.5$  for traits with low heritability (Knapp, 1994).

It should be emphasized that a two-stage selection procedure or independent culling does not increase the total aggregate genetic gain *per se*; rather, it increases the gain per unit cost. The results of this



study show that the two-stage MIS with respect to the gain per unit cost is inferior to conventional phenotypic selection when  $h^2 \geq 0.3$  and the cost ratio ( $r$ ) of obtaining measurements on phenotypic characters to scoring marker loci is smaller than unity. The efficiency of multistage MIS to conventional selection increases as  $r$  increases and  $h^2$  decreases. In some species with a long life cycle or a large body volume such as beef cattle and trees, it is quite possible that  $r > 1.0$ , i.e. the cost of obtaining measurements on phenotypic characters is more expensive than that of scoring marker loci. In these situations, multistage MIS with respect to the gain per unit cost would be expected to be more effective than conventional index selection. In view of this, the initial applications of MAS in commercial breeding will be in those situations where MAS is likely to provide additional advantages, such as in the case of sex-limited traits, for characters that are difficult or expensive to record, or in species where individuals have high value.

### Acknowledgements

We thank Damian Gessler for his helpful comments. This research was supported by NRI Competitive Grants Program/USDA 95-37205-2313 and NIH Grant GM55321-01 to S.X.

### References

- CUNNINGHAM, E. P. 1975. Multi-stage index selection. *Theor. Appl. Genet.*, **46**, 55-61.
- DUDLEY, J. W. 1993. Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. *Crop Sci.*, **33**, 660-668.
- EDWARDS, M. D. AND PAGE, N. J. 1994. Evaluation of marker-assisted selection through computer simulation. *Theor. Appl. Genet.*, **88**, 376-382.
- EDWARDS, M. D., STUBER, C. W. AND WENDEL, J. F. 1987. Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, distribution, and types of gene action. *Genetics*, **116**, 113-125.
- FEHR, W. R. 1987. *Principles of Cultivar Development*, vol. 1, *Theory and Technique*. Macmillan, New York.
- FERNANDO, R. L. AND GROSSMAN, M. 1989. Marker assisted selection using best linear unbiased prediction. *Génét. Sél. Évol.*, **21**, 467-477.
- GIMELFARB, A. AND LANDE, R. 1994. Simulation of marker assisted selection in hybrid populations. *Genet. Res.*, **63**, 39-47.
- HALEY, C. S. AND KNOTT, S. A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, **69**, 315-324.
- HELENTJARIS, T., SLOCUM, T. M., WRIGHT, S., SCHAEFER, A. AND NIENHUIS, J. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor. Appl. Genet.*, **72**, 761-769.
- JIANG, C. AND ZENG, Z.-B. 1995. Multiple-trait analysis of genetic mapping for quantitative trait loci. *Genetics*, **140**, 1111-1127.
- KNAPP, J. S. 1994. Selection using molecular marker indexes. In: *Analysis of Molecular Marker Data*, Proc. Joint Symp. Am. Soc. Hort. Sci./Crop Sci. Soc. Am. 5-6 Aug 1994, pp. 3-7. Corvallis, OR.
- LANDE, R. 1992. Marker-assisted selection in relation to traditional methods of plant breeding. In: Stalker, H. T. and Murphy, J. P. (eds) *Plant Breeding in the 1990s*, pp. 437-451. C.A.B. International, Wallingford.
- LANDE, R. AND THOMPSON, R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, **124**, 743-756.
- MEUWISSEN, T. H. E. AND VAN ARENDONK, J. A. M. 1992. Potential improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. *J. Dairy Sci.*, **75**, 1651-1659.
- NAMKOONG, G. 1970. Optimal allocation of selection intensity in two stages of truncation selections. *Biometrics*, **26**, 465-476.
- NEIMANN-SORENSEN, A. AND ROBERTSON, A. 1961. The association between blood groups and several production characters in three Danish cattle breeds. *Acta Agric. Scand.*, **11**, 163-196.
- RUANE, J. AND COLLEAU, J. J. 1995. Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genet. Res.*, **66**, 71-83.
- SMITH, C. 1967. Improvement of metric traits through specific genetic loci. *Anim. Prod.*, **9**, 349-358.
- SOLLER, M. AND BECKMANN, J. S. 1983. Genetic polymorphism in variety identification and genetic improvement. *Theor. Appl. Genet.*, **67**, 25-33.
- STRAUSS, S. H., LANDE, R. AND NAMKOONG, G. 1992. Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. Forest Res.*, **22**, 1050-1061.
- STUBER, C. W. 1992. Biochemical and molecular markers in plant breeding. *Pl. Breed. Res.*, **9**, 37-61.
- STUBER, C. W., EDWARDS, M. D. AND WENDEL, J. F. 1987. Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. *Crop Sci.*, **27**, 639-648.
- TANKSLEY, S. D., MEDINO-FILHO, H. AND RICK, C. M. 1981. The effect of isozyme selection on metric characters in an interspecific backcross of tomato: Basis of an early screen procedure. *Theor. Appl. Genet.*, **60**, 291-296.
- XIE, C. AND XU, S. 1996. Best linear unbiased prediction under selection. *J. Genet. Breed.*, **50**, 287-294.
- XU, S. AND MUIR, W. M. 1991. Multistage selection for genetic gain by orthogonal transformation. *Genetics*, **129**, 963-974.
- XU, S. AND MUIR, W. M. 1992. Selection index updating. *Theor. Appl. Genet.*, **83**, 451-458.
- ZHANG, W. AND SMITH, C. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theor. Appl. Genet.*, **83**, 813-820.

**Appendix A**

*Derivation of eqns (9) and (10)*

Let  $\mathbf{x}$  be a vector of quantitative traits with phenotypic and additive genetic variance-covariance matrices  $\text{Var}(\mathbf{x}) = \mathbf{P}$  and  $\text{Var}(\mathbf{g}) = \mathbf{G}$ ,  $\mathbf{m}$  is a vector of corresponding molecular scores obtained by summing the vectors of effects on the traits produced by associated molecular marker loci, and  $\text{Var}(\mathbf{m}) = \text{Cov}(\mathbf{g}^T, \mathbf{m}) = \mathbf{M}$ . Then the full data vector is  $\mathbf{y} = (\mathbf{m} \ \mathbf{x})^T$ . These have relative economic weights  $\bar{\mathbf{w}} = (\mathbf{0} \ \mathbf{w})^T$  and the phenotypic ( $\mathbf{V}$ ) and genotypic ( $\mathbf{A}$ ) variance-covariance matrices

$$\mathbf{V} = \begin{pmatrix} \mathbf{M} & \mathbf{M} \\ \mathbf{M} & \mathbf{P} \end{pmatrix} \text{ and } \mathbf{A} = \begin{pmatrix} \mathbf{M} & \mathbf{M} \\ \mathbf{M} & \mathbf{G} \end{pmatrix},$$

respectively. Let  $\mathbf{z} = [z_1 \ z_2]^T$  be a  $2 \times 1$  vector of the updated selection indices defined by Xu & Muir (1992), i.e.:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{11}^T & \mathbf{0} \\ \mathbf{b}_{12}^T & \mathbf{b}_{22}^T \end{pmatrix} \begin{pmatrix} \mathbf{m} \\ \mathbf{x} \end{pmatrix} = \mathbf{B}^T \mathbf{y}. \tag{A1}$$

The  $\mathbf{b}_{ij}$ s are found so that  $\text{Cov}(H, z_i)$  is maximum, or equivalently  $E[(\mathbf{b}_i^T \mathbf{y} - \bar{\mathbf{w}}^T \mathbf{g})^T (\mathbf{b}_i^T \mathbf{y} - \bar{\mathbf{w}}^T \mathbf{g})]$  is minimum subject to the restriction  $\text{Cov}(z_1, z_2) = \mathbf{b}_2^T \mathbf{V}_{2,1} \mathbf{b}_1 = 0$ , where  $\mathbf{V}_{2,1}$  is a submatrix with two rows and the first column of  $\mathbf{V}$ . Obviously, this procedure starts at the second stage.

At the first stage, selection is based on  $\mathbf{m}$  only. We have  $z_1 = \mathbf{b}_{11}^T \mathbf{m}$  where  $\mathbf{b}_{11} = \mathbf{M}^{-1}(\mathbf{M} \ \mathbf{M}) \bar{\mathbf{w}} = \mathbf{w}$ . At the second stage, selection is based on both  $\mathbf{x}$  and  $\mathbf{m}$ . The index coefficients for the second stage have the form of  $\mathbf{b}_2 = (\mathbf{b}_{12} \ \mathbf{b}_{22})^T$ . After introducing Lagrange multipliers of  $\tau$ , where  $\tau$  is an  $i-1$  vector, the optimum  $\mathbf{b}_2$  is found by minimizing the following quantity with respect to  $\mathbf{b}_2$ :

$$\mathbf{T} = \mathbf{b}_2^T \mathbf{V} \mathbf{b}_2 - 2 \mathbf{b}_2^T \mathbf{A} \bar{\mathbf{w}} + \bar{\mathbf{w}}^T \mathbf{A} \bar{\mathbf{w}} - 2 \tau^T (\mathbf{b}_2^T \mathbf{V}_{2,1} \mathbf{b}_1). \tag{A2}$$

After vector differentiation and simplification, we have:

$$\mathbf{b}_2 = [\mathbf{I} - \mathbf{V}^{-1} \mathbf{V}_{2,1} \mathbf{b}_1 (\mathbf{b}_1^T \mathbf{V}_{1,2} \mathbf{V}^{-1} \mathbf{V}_{2,1} \mathbf{b}_1)^{-1} \mathbf{b}_1^T \mathbf{V}_{1,2}] \mathbf{V}^{-1} \mathbf{A} \bar{\mathbf{w}}, \tag{A3}$$

where  $\mathbf{I}$  is an identity matrix with the same dimension as  $\mathbf{V}$ . Equivalently,  $\mathbf{b}_2$  can be expressed as:

$$\mathbf{b}_2 = [-(\mathbf{P} - \mathbf{M})^{-1} (\mathbf{G} - \mathbf{M}) \mathbf{w} (\mathbf{P} - \mathbf{M})^{-1} (\mathbf{G} - \mathbf{M}) \mathbf{w}]^T.$$

Let  $\mathbf{b}_1 = (\mathbf{w} \ \mathbf{0})^T$  and  $\mathbf{B} = (\mathbf{b}_1 \ \mathbf{b}_2)$ ; this yields text eqn (9).

The standard normal  $\mathbf{z}$  can be achieved by standardizing the  $\mathbf{b}_i$ s,  ${}^* \mathbf{B} = \mathbf{B}(\mathbf{B}^T \mathbf{P} \mathbf{B})^{-1/2}$ . So that  ${}^* \mathbf{B} \mathbf{P} {}^* \mathbf{B} = \mathbf{I}$ , where  $\mathbf{I}$  is an identity matrix. The matrix  $\mathbf{B}$  after standardization is:

$$\mathbf{B} = {}^* \mathbf{B} = \begin{pmatrix} 1/\mathbf{M}^{1/2} & -(\mathbf{P} - \mathbf{M})^{-1/2} \\ 0 & (\mathbf{P} - \mathbf{M})^{-1/2} \end{pmatrix}. \tag{A4}$$

The vector of genetic gain,  $\Delta \mathbf{G}$ , is predicted by  $\Delta \mathbf{G} = \mathbf{A}^T \mathbf{B} \Delta \mathbf{z}$ , which gives eqn (10) in the text.

By generalizing the procedure, the two-stage selection is readily extended to multistage selection. Let  $\mathbf{b}_i$  be the index coefficients for the  $i$ th stage selection, then eqn (A3) can be modified as:

$$\mathbf{b}_i = [\mathbf{I} - \mathbf{V}_{ii}^{-1} \mathbf{V}_{i(i-1)} \mathbf{B}_{(i-1)} (\mathbf{B}_{(i-1)}^T \mathbf{V}_{(i-1)} \mathbf{V}_{ii}^{-1} \mathbf{V}_{i(i-1)} \mathbf{B}_{(i-1)})^{-1} \mathbf{B}_{(i-1)}^T \mathbf{V}_{(i-1)}] \mathbf{V}_{ii}^{-1} \mathbf{A}_i \bar{\mathbf{w}}, \tag{A5}$$

where  $\mathbf{V}_{ij}$ ,  $\mathbf{B}_i$  and  $\mathbf{A}_i$  are submatrices of  $\mathbf{V}$ ,  $\mathbf{B}$  and  $\mathbf{A}$  corresponding to the  $i$ th stage selection.

**Appendix B**

*The values of matrices P, G and M for the numerical examples*

In the case of two-trait selection, the values of heritability ( $h^2$ ), genotypic ( $r_g$ ) or phenotypic ( $r_p$ ) correlations, and the proportion of additive genetic variance associated with marker loci ( $p$ ) used for alternative selection procedures are obtained as follows. Assuming that each trait has a unit phenotypic variance, we have:

$$\mathbf{P} = \begin{pmatrix} 1 & r_p \\ r_p & 1 \end{pmatrix} \text{ and } \mathbf{G} = \begin{pmatrix} h_1^2 & t \\ t & h_2^2 \end{pmatrix},$$

where  $t = r_g \sqrt{h_1^2 h_2^2}$ . Because  $\mathbf{M} \mathbf{G}^{-1}$  is the multiple-trait form of the proportions of additive genetic variances associated with marker effects, let:

$$\mathbf{M} \mathbf{G}^{-1} = \begin{pmatrix} p_1 h_1^2 & t' \\ t' & p_2 h_2^2 \end{pmatrix},$$

where  $0 \leq p_1, p_2 \leq 1$ ,  $t' = p_1 p_2 t$  when  $h_1^2 = h_2^2$  and  $p_1 = p_2$ , and

$$t' = \frac{(p_2 h_2^2 - p_1 h_1^2) t}{h_2^2 - h_1^2}$$

when  $h_1^2 \neq h_2^2$ . Then,  $\mathbf{M} = \mathbf{M} \mathbf{G}^{-1} \mathbf{G}$ , where  $\mathbf{M}$  is a symmetric matrix. In practice, the values of  $\mathbf{M}$  are not known. But here we are interested in the proportions of MQTL effects to the total additive genetic variance for each trait.

Similarly, in the case of four-trait selection, assuming that  $r_{pi} = r_{pj}$ ,  $r_{gi} = r_{gj}$ ,  $p_i = p_j$ , and  $h_i^2 = h_j^2$  for  $i, j = 1, \dots, 4$  and each trait has a unit phenotypic variance, we have  $t = r_g h_i^2$  and  $t' = p_i^2 t$ .