*Genet. Res., Camb.* (1994), **64**, *pp.* 127–136 *With 5 text-figures*  

# Simulation of marker assisted selection for non-additive traits

A. GIMELFARB AND R. LANDE*
*Department of Biology, University of Oregon, Eugene, Oregon 97403, U.S.A.*

(*Received February 24, 1994 and in revised form June 22, 1994*)

**Summary**

Marker Assisted Selection (MAS) based on additive effects associated with alleles at marker loci, estimated by linear regression of individual phenotype on the markers, was applied to characters with non-additive gene action and non-additive environment. The base population was the $F_2$ generation of a cross between two inbred lines. Computer simulations of MAS were conducted for characters with dominance, epistasis and genotype-environment interaction approximated by the 'additive-multiplicative' model. MAS was more effective than purely phenotypic selection in all cases. The efficiency of MAS for characters with non-additive gene action is comparable to (and for negative dominance even higher than) the efficiency of MAS for strictly additive characters. Environmental non-additivity, however, lowers the efficiency of MAS. Almost all results concerning the efficiency of MAS in our previous simulations of purely additive traits are applicable to non-additive traits.

## 1. Introduction

A method of Marker Assisted Selection (MAS) proposed by Lande & Thompson (1990) utilizes the linkage disequilibrium between molecular genetic marker loci and quantitative trait loci (QTLs) in populations created by a cross between two inbred lines. Additive effects of marker genes on the trait are estimated by multiple linear regression of the phenotype on the markers, and selection is based on an index incorporating the phenotype and the estimated effects of markers. Gimelfarb & Lande (1994) reported results of computer simulations of MAS using this method. The simulations confirmed the conclusions reached by Lande & Thompson from their theoretical analysis that selection for a quantitative character based on the estimated additive effects of markers is generally more effective (especially in the initial generations) than selection based strictly on phenotypes. Zhang & Smith (1992, 1993) conducted similar simulations of MAS. They, however, compared MAS not to purely phenotypic selection but rather to selection based on the BLUP estimate of an individual's breeding value. Their conclusion was that selection based on an index combining the effects of markers and the BLUP estimate is more effective than selection based solely on the BLUP estimate.

* Corresponding Author.

Characters in simulations by Zhang & Smith (1992, 1993) and by Gimelfarb & Lande (1994) were additive, i.e., genes as well as environment contributed additively to the phenotype. This is important since, even though the method of Lande & Thompson (1990) can be applied to non-additive characters, their theoretical analysis of the efficiency of MAS required the additivity assumption. No theoretical work or computer simulation has been done so far on the effectiveness of MAS for characters that are not additive. It is important, however, to know whether MAS can be more effective than purely phenotypic selection for such characters as well. To answer this question we conducted a series of computer simulations of MAS based on the method by Lande & Thompson, but with characters that are not additive.

We investigated the following deviations from the strictly additive model of quantitative traits: non-additivity within a locus (dominance), non-additivity between loci (epistasis), and non-additivity of the environmental effect (genotype-environment interaction, $G \times E$). There are many biological mechanisms that may cause these non-additivities, and, almost certainly, different mechanisms are responsible for non-additivity of different traits in different organisms. Our goal, however, was not to investigate a particular biological mechanism, but rather to find out whether MAS based on the additive effects of marker genes estimated by linear regression can be more effective

than selection based strictly on the phenotype, even if the additivity assumption does not hold. To this end we have employed an 'additive-multiplicative' approximation for dominant and epistatic gene interactions (Gimelfarb, 1989) as well as for the interaction between genotype and environment.

## 2. Methods

The genetic map used in the simulations consisted of 10 chromosomes of 100 cM each. All chromosomes carried marker genes equidistantly spaced along the chromosome. The number of markers was the same for each chromosome. Six out of ten chromosomes also carried a total of 10 QTLs distributed among them as shown in Fig. 1, in which a square indicates a QTL. Also shown there are genetic markers (11 per chromosome). The remaining four chromosomes did not have QTLs on them (but did carry marker genes). Each QTL had two alleles, 0 and 1, and all loci were equivalent in their effect on the character.

The effect by the $i$th QTL in the individual's genotype on the character was calculated as

$$z_i = (1+D)(Q_i' + Q_i'') - 2DQ_i' Q_i'', \tag{1}$$

where $Q_i'$ and $Q_i''$ are the homologous alleles (0 or 1) at the locus. Parameter $D$ represents the degree of dominance. Indeed, it is not difficult to see that (1) yields the following effect by a QTL:

$$\left.\begin{array}{lll} z_i = 0 & \text{in} & 0/0 \text{ homozygote}, \\ z_i = 1+D & \text{in} & 0/1 \text{ heterozygote}, \\ z_i = 2 & \text{in} & 1/1 \text{ homozygote}. \end{array}\right\} \tag{2}$$

Hence, allele 1 is dominant if $D > 0$, allele 0 is dominant if $D < 0$ and allelic contributions are additive if $D = 0$.

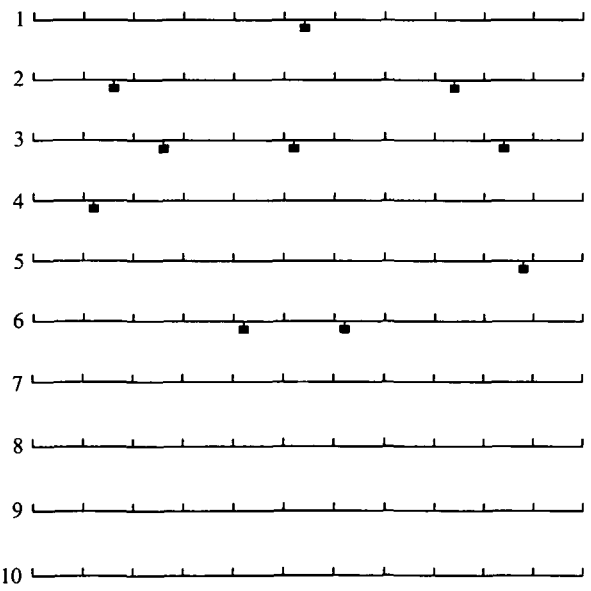


Fig. 1. Genetic map of 10 chromosomes with 11 markers per chromosome and a total of 10 QTLs of equal effect (indicated by squares).

Given the effects on the character by all QTLs in the genotype of an individual, the individual's genotypic value was computed as

$$z = (1-C)\sum_i z_i + C\prod_i z_i, \tag{3}$$

where the parameter $0 \leqslant C \leqslant 1$ determines the strength of the multiplicative interaction between loci. If $C = 0$, the contributions by QTLs to the genotypic value are strictly additive, whereas they are strictly multiplicative if $C = 1$. Thus, the model (3) can be viewed as a combination of two scales widely used in practical applications: linear and logarithmic.

Given the genotypic value of an individual, $z$, its phenotypic value was determined as

$$Z' = (1-B)(z+e) + Bze, \tag{4}$$

where $e$ is a random variable representing the environmental effect on the character which is assumed normally distributed with zero mean. Thus, instead of being just the sum of the genotypic value and the environmental effect, as in the traditional additive model, the phenotype in our model is a mixture of their sum and their product. If $B = 0$, the environmental contribution is strictly additive, whereas it is strictly multiplicative if $B = 1$. It should be pointed out that when discussing environment in this paper we refer only to microenvironment, i.e., non-genetic differences between individuals within a population. We do not consider macroenvironment, i.e., non-genetic differences between populations. The variance of the environmental effect, $v_e$, was computed at the beginning of each run, and it remained unchanged in subsequent generations of the run. For purely additive environment, the variance $v_e$ was calculated in the same way as in our previous paper, whereas the calculation of $v_e$ for additive-multiplicative environment was based on equation (A 13) in the Appendix. The phenotype of an individual was scaled by the following linear transformation of the value determined by formula (4):

$$Z = 20[Z'/(1-B) - z_{\min}]/(z_{\max} - z_{\min}) - 10, \tag{5}$$

where $z_{\min}$ and $z_{max}$ are the minimum and the maximum of the genotypic values among all possible QTL genotypes (we have considered only environments with $B < 1$). Given that the mean of the environment effect is zero, transformation (5) constrains the mean phenotype in a population to the interval between $-10$ and 10. Consequently, the mean value of any trait under any selection could not exceed 10.

Except for the processes determining the phenotype, all other processes were simulated as in our previous paper (Gimelfarb & Lande, 1994). Selection was based on the index,

$$I = Z + b_M M, \tag{6}$$

incorporating the phenotype, $Z$, as well as the molecular score, $M$, of an individual. The molecular score for an individual was equal to the sum of the

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

Table 1. *Parameters used in simulations*

| | BASE | ALTERNATE |
|---|---|---|
| Markers on chromosome | 11 | 3 6 21 51 |
| Markers in selection index | 6 | 3 9 12 15 |
| Individuals of each sex | 500 | 100 200 3000 |
| Initial heritability | 0·1 | |
| Selection strength | 25% | |

additive effects associated with alleles at marker loci included in the index (the coefficients of a linear regression of the phenotype on the markers). The molecular score coefficient, $b_M$, was calculated as

$$b_M = (1/H^2 - 1)/(1 - p_M), \quad (7)$$

where $H^2$ is the broad heritability and $p_M$ is the proportion of the genotypic variance in the population accounted for by the markers included in the index. The heritability in a given generation was computed as the ratio of the variance of the genotypic values (calculated based on expression (3)) to the phenotypic variance. Parameter $p_M$ was computed as

$$p_M = R^2/H^2, \quad (8)$$

where $R^2$ is the coefficient of determination for the linear regression. Lande & Thompson (1990) and Gimelfarb & Lande (1994) used the narrow heritability rather than the broad heritability in computing $b_M$ and $p_M$. Given, however, that both papers dealt with purely additive traits, the two heritabilities were the same. A two-stage multiple regression procedure described in our previous paper was used to choose marker loci for inclusion in the index out of the total number of markers in the genome and to estimate 'additive effects' associated with the chosen marker loci. This procedure was employed in each generation of a computer run. No correction was made for the bias in $R^2$ due to the non-random selection of markers in the regression procedure, since our simulations for additive characters indicate that the effect of the correction is negligible if the population size is above 100 individuals of each sex.

Parameters used in the computer simulations for non-additive traits are shown in Table 1. They are similar to those used in our simulations for additive traits (Gimelfarb & Lande, 1994). The majority of simulations were conducted with parameters in the BASE set, but other parameters were also used. In such cases, only one ALTERNATE parameter was substituted in the BASE set at a time. Each run was initiated from the $F_2$ generation of a cross between two inbred lines, so that the initial allelic frequencies were 1/2 at all QTLs and marker loci. Two distinct types of initial populations with respect to the gametic phase of QTLs were considered: total coupling (alleles of equal effect in adjacent QTLs on a chromosome) and total repulsion (alleles of opposite effect at adjacent QTLs on a chromosome). Simulations for a given set of parameters and an initial gametic phase were replicated, and a result reported here represents the average over 40 runs for populations of 100 and 200 individuals of each sex, over 30 runs for 500 individuals of each sex, and over 20 runs for 3000 individuals of each sex.

## 3. Results and Discussion

Only one deviation from the strictly additive model of quantitative characters was investigated at a time. Thus, neither epistasis nor non-additive environment were present in runs with dominance. On the other hand, there was no dominance and the environment was additive in runs with epistasis. A range of parameters $D$ for dominance, $C$ for epistasis and $B$ for non-additive environment were investigated, but results for only some of them are reported here, namely, $D = 1$ (complete positive dominance), $D = -1$ (complete negative dominance), epistasis with $C = 0{\cdot}01$, and 'additive-multiplicative' environment with $B = 0{\cdot}1$, $0{\cdot}5$, $0{\cdot}8$.

Table 2 presents the variance components (as proportions of the total phenotypic variance) in the initial population. Their computation is explained in the Appendix. Depending on the initial gametic phase, the non-additive component of variance accounts for 20% or 50% of the total genotypic variance in the case of positive dominance, and for 40% or 50% in the case of negative dominance. On the other hand, the non-additive component in the first generation is practically zero for characters with epistasis. This does not mean, however, that there is practically no interaction between the contributions by the QTLs to the character. Indeed, the non-additive component of variance is determined not only by the mode and strength of gene interaction but also by the distribution of the genotypes in the population, and it increases in later generations of selection. The variance component due to genotype-environment interaction for characters with additive-multiplicative environment

Table 2. *Variance components (as proportions of the total phenotypic variance) in the initial* $F_2$ *population.*

| | | $V_A$ | $V_{NA}$ | $V_E$ | $V_{GE}$ |
|---|---|---|---|---|---|
| Dominance | C | 0·08 | 0·02 | 0·90 | 0·00 |
| (+) | R | 0·05 | 0·05 | 0·90 | 0·00 |
| Dominance | C | 0·06 | 0·04 | 0·90 | 0·00 |
| (−) | R | 0·05 | 0·05 | 0·90 | 0·00 |
| Epistasis | C | 0·10 | 0·00 | 0·90 | 0·00 |
| | R | 0·10 | 0·00 | 0·90 | 0·00 |
| $G \times E$ | C | 0·10 | 0·00 | 0·82 | 0·08 |
| (0·1) | R | 0·10 | 0·00 | 0·87 | 0·03 |
| $G \times E$ | C | 0·10 | 0·00 | 0·10 | 0·80 |
| (0·5) | R | 0·10 | 0·00 | 0·23 | 0·67 |
| $G \times E$ | C | 0·10 | 0·00 | 0·01 | 0·89 |
| (0·8) | R | 0·10 | 0·00 | 0·02 | 0·88 |

($V_A$ – additive genetic; $V_{NA}$ – non-additive genetic; $V_E$ – environmental; $V_{GE}$ – genotype-environment interaction; C and R indicate total coupling and total repulsion initial gametic phase).

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

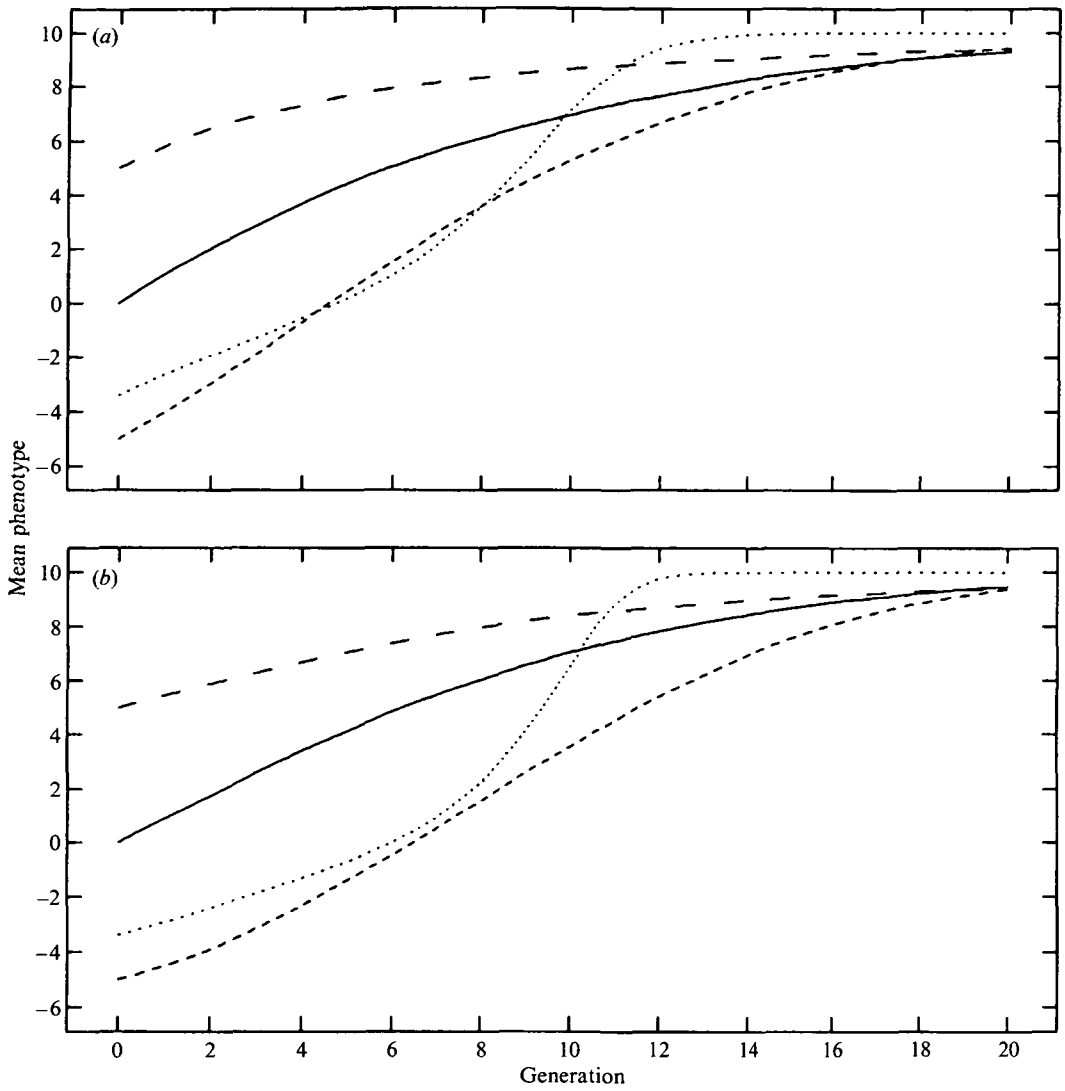


Fig. 2. Dynamics of the mean phenotype of characters with non-additive gene action under purely phenotypic selection (*a*) coupling; (*b*) repulsion. ....., epistasis; ----, dominance (−); ———, dominance (+); ——, additivity.

changes from only 8% or 3% (depending on the initial gametic phase) of the total phenotypic variance if the coefficient $B$ in (4) is 0·1 to almost 90% if $B =$ 0·8.

Figure 2 shows the dynamics of the mean phenotype of characters with additive and non-additive gene action under purely phenotypic selection for higher values of the trait. Both dominance and epistasis are seen to have a noticeable effect on the dynamics. Characters with negative dominance show the largest overall gain in the mean after 20 generations of selection, whereas positive dominance yields the smallest gain. The response by characters with epistasis is relatively slow initially, but after generation 6 it accelerates appreciably, reaching a plateau by generation 12, while traits with additive and dominant gene action continue responding until generation 20. Hence, non-additivity in gene action may influence the response to selection. The accelerated response by characters with epistasis was accompanied by a rise in the broad sense heritability. Starting from the initial value of 0·1, the broad heritability was declining initially, but then went up, reaching at generation 9 a maximum of 0·22 if the initial gametic phase was coupling, and a maximum of 0·39 if the gametic phase was repulsion.

The effectiveness of Marker Assisted Selection relative to purely phenotypic selection for characters with additive or non-additive gene action can be seen in Fig. 3 which shows the efficiencies of MAS with the BASE parameter set. Efficiency is defined as the ratio of the response by the mean phenotype under MAS to the response under purely phenotypic selection with the same set of parameters. Efficiency higher (lower) than 1 means that MAS is more (less) effective than phenotypic selection. Fig. 3 clearly indicates that MAS based on additive effects associated with marker loci estimated by a linear regression is effective not only for purely additive characters but for characters with non-additive gene action as well. Moreover, for characters with negative dominance, the efficiency is much higher than for additive characters. The efficiency for characters with positive dominance is generally lower than for additive characters (except

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

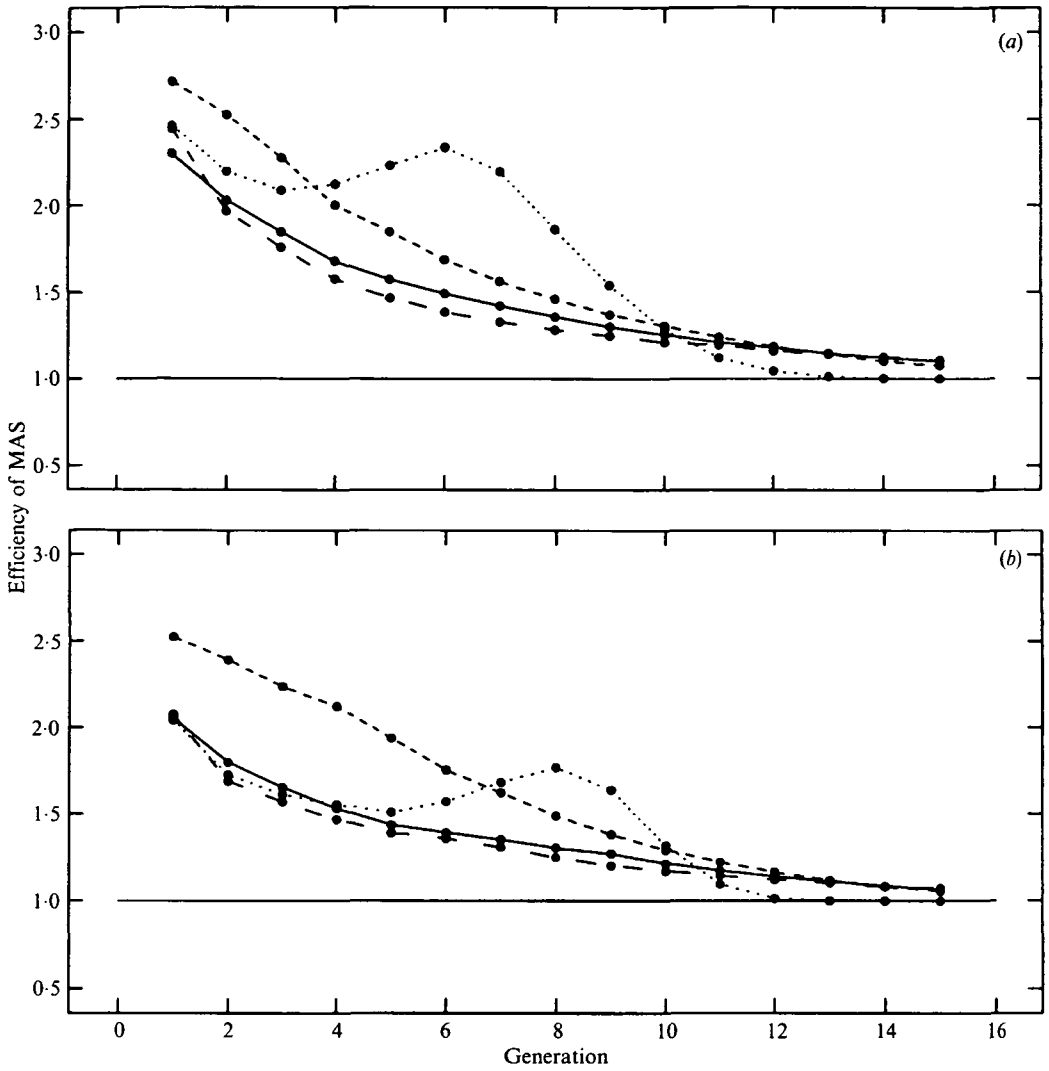


Fig. 3. Efficiency of MAS relative to purely phenotypic selection for characters with non-additive gene action (*a*) coupling; (*b*) repulsion. ....., epistasis; – – – –, dominance (−); ————, dominance (+); ——, additivity.

generation 1 in coupling), but the difference is quite small.

Quite unexpected is the behaviour of the efficiency of MAS for characters with multiplicative epistasis. Indeed, given that MAS utilizes linkage disequilibrium between marker loci and QTLs, and that this disequilibrium is reduced by recombination, the efficiency of MAS is expected to decline with each generation of selection. This was shown in our previous paper to be true for additive characters, and it is also true for characters with dominance (Fig. 3). The same is not true, however, for characters with multiplicative epistasis (Fig. 3). In this case, efficiency does not go down in each generation, but rather has a 'hump', i.e., after an initial decline, efficiency starts climbing up, reaches a maximum at generation 6 or 8 (depending on the initial gametic phase) and then goes down again. The efficiency on the top of the 'hump' is quite high, almost comparable to that in the first generation, especially for coupling gametic phase. This behaviour of the efficiency is accompanied by a similar behaviour of the broad heritability. Starting from an initial value 0·1, the broad heritability also declines initially but then rises, reaching a maximum of 0·20 at generation 6 (coupling) or a maximum of 0·38 at generation 8 (repulsion), i.e., at the same generation at which the 'hump' of the efficiency has its peak (Fig. 3).

Besides investigating multiplicative epistasis, we have also conducted a few simulations of MAS using the formula

$$z = \sum_i z_i + C \sum_i \sum_j z_i z_j \quad (i \neq j) \tag{9}$$

instead of formula (3) to compute the genotypic value of an individual. In simulations with the BASE parameter set, the efficiency of MAS for characters with strong 'pairwise' epistasis ($C = 1$ in the above formula) was similar to the efficiency of MAS for purely additive characters. The trajectory of the efficiency did not have a 'hump' as in the case of multiplicative epistasis, but declined monotonically.

Let us now turn to characters with purely additive gene action but with 'additive-multiplicative' en-

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

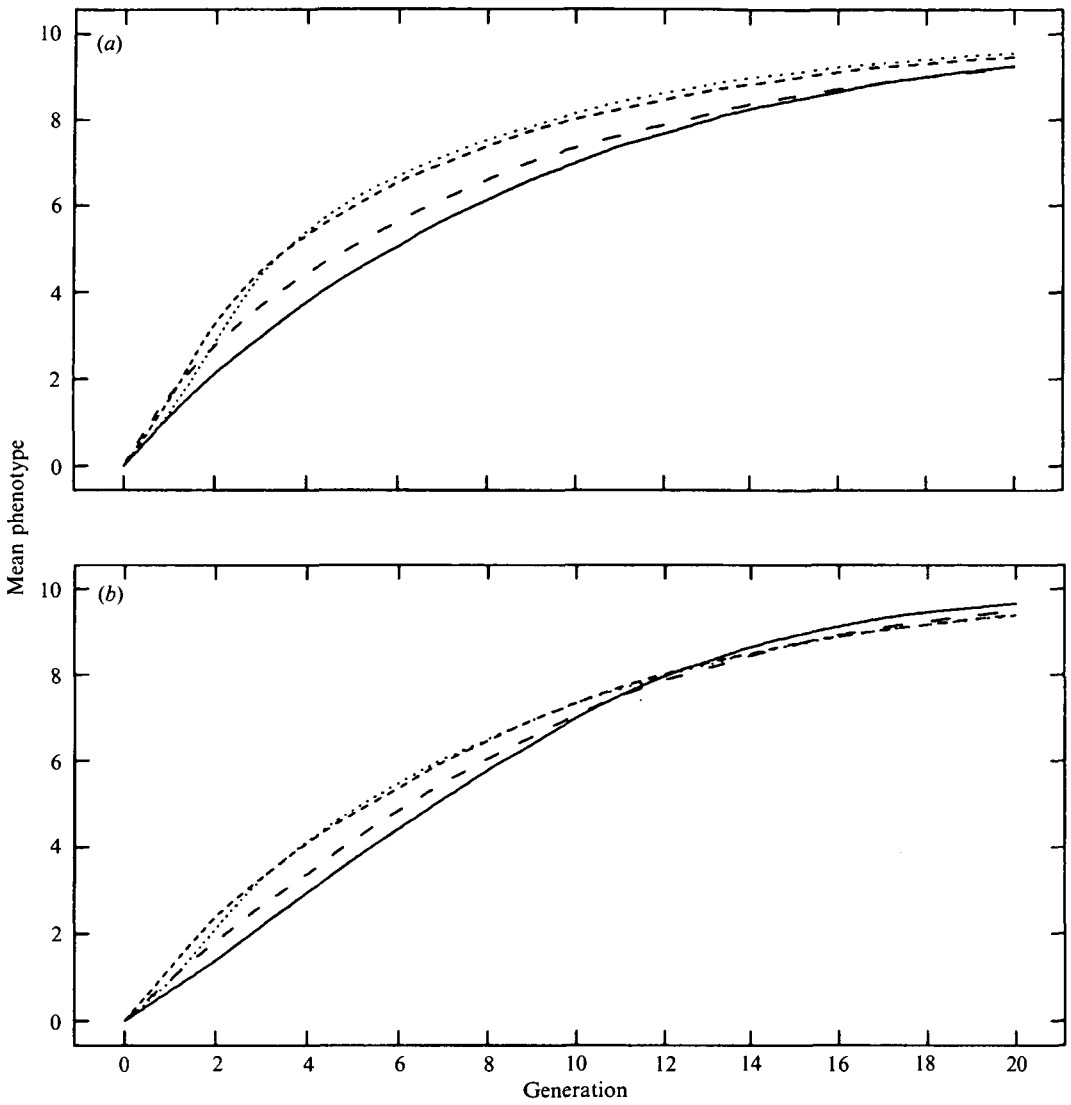


Fig. 4. Dynamics of the mean phenotype of characters with non-additive environment under purely phenotypic selection (*a*) coupling; (*b*) repulsion. ....., $G \times E$ (0·8); ----, $G \times E$ (0·5); ———, $G \times E$ (0·1); ——, additivity.

vironment. Figure 4 shows the response by such characters to purely phenotypic selection. It is seen that, while the effect of non-additivity in gene action on the response to purely phenotypic selection is quite substantial (Fig. 2), the effect of environmental non-additivity is not very large, especially if the initial gametic phase is repulsion, although it is more noticeable for coupling gametic phase. In general, characters with additive-multiplicative environment respond to phenotypic selection faster than characters with purely additive environment. Notice also that the response in the first generation with $B = 0{\cdot}8$ (high non-additivity) is slower than the response with $B = 0{\cdot}1$ (low non-additivity). After generation 2, however, the response with $B = 0{\cdot}8$ becomes faster than with $B = 0{\cdot}1$.

Figure 5 shows the efficiency of MAS with the BASE parameter set for characters with purely additive or additive-multiplicative environment. It is seen that the influence of environmental non-additivity on the efficiency of MAS is noticeably greater than on the response to purely phenotypic selection (Fig. 4). The efficiency is generally much higher if the environment is additive than if it is non-additive. This is because in the presence of genotype-environment interaction (even in an infinite population), the additive effects associated with markers include not only a contribution from QTLs but also a contribution from the environment. Thus, even though the additive effects associated with markers may be statistically significant and may predict well the phenotypes in the same generation in which they have been evaluated, they are not as good at predicting phenotypes in the following generation because the genotype-environment interaction changes during evolution. The lowest efficiency is for MAS in the case of non-additive environment with the coefficient $B$ in (4) equal to 0·5. In fact, after two or three generations (depending on the gametic phase) MAS for such a character becomes less effective than purely phenotypic selection. With $B = 0{\cdot}8$ MAS also becomes less effective than phenotypic selection after two or three generations. In the first

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

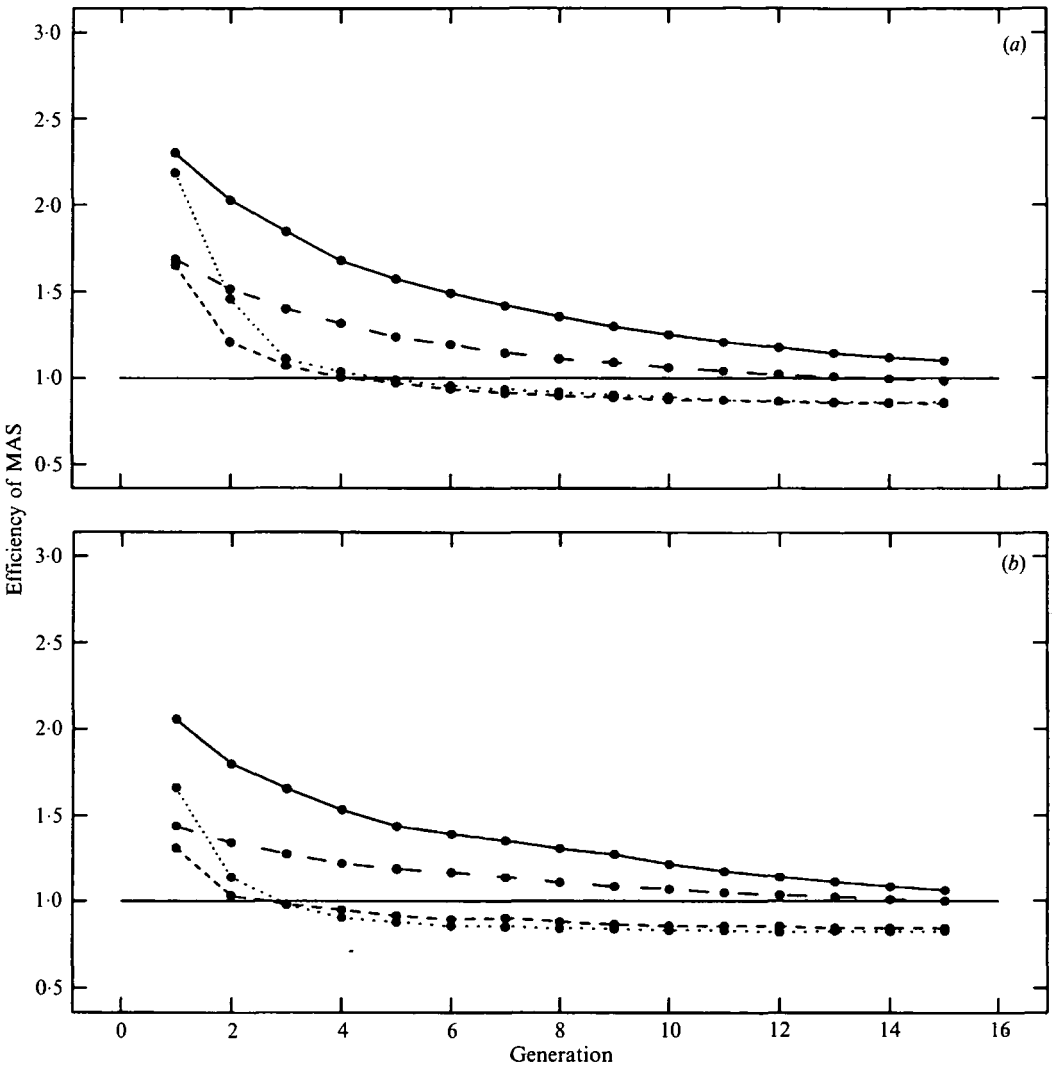


Fig. 5. Efficiency of MAS relative to purely phenotypic selection for characters with non-additive environment (*a*) coupling; (*b*) repulsion. ....., $G \times E$ (0·8); ----, $G \times E$ (0·5); ———, $G \times E$ (0·1); ——, additivity.

generation, however, the efficiency with $B = 0{\cdot}8$ is higher than for characters with $B = 0{\cdot}5$ or $B = 0{\cdot}1$. Moreover, if the gametic phase is coupling, the efficiency in the case of $B = 0{\cdot}8$ nearly equals the efficiency with strictly additive environment.

We demonstrated in our previous paper that, while the response by an additive character to purely phenotypic selection is practically not affected by changes in population size above 100 individuals of each sex, population size is the most important factor influencing the efficiency of MAS for additive characters. The same is also true for characters with non-additive gene action and with non-additive environment. Their response to purely phenotypic selection is almost the same for any population size above 100 individuals of each sex, whereas the efficiency of MAS is always higher in a population of a larger size. Table 3 demonstrates the influence of the population size on the efficiency in the first generation of MAS for characters with different types of non-additivity. It is seen that the influence can be very strong. For example, changing the population size from 100 to 500 individuals of each sex increases efficiency of MAS by between 70% and 100% for dominance or epistasis and between 40% and 50% for non-additive environment.

Runs with ALTERNATE parameters (Table 1) yielded results similar to those reported by (Gimelfarb and Lande, 1994) for additive traits. Increasing the number of marker loci on a chromosome, or including more markers in the selection index does not necessarily result in an increased efficiency of MAS. In fact, there is an optimum number of markers per chromosome as well as an optimum number of markers included in the index that give the highest response to MAS. The exact value of the optimum depends on other parameters as well as on the initial gametic phase. For example, given 11 markers per chromosome, the optimum number of markers for inclusion in the index was 6 in almost all simulations, except for those with epistasis and repulsion gametic phase, for which the highest response to MAS was if 15 rather than six markers were included in the index. In general, however, changing the number of markers on a

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

Table 3. *The efficiency in the first generation of MAS in populations of different sizes*

| | | Population size | | | |
|---|---|---|---|---|---|
| | | 100 | 200 | 500 | 3000 |
| Dominance (+) | C | 1·41 | 1·93 | 2·44 | 2·87 |
| | R | 0·94 | 1·53 | 2·08 | 3·09 |
| Dominance (−) | C | 1·45 | 2·16 | 2·71 | 3·40 |
| | R | 1·45 | 1·74 | 2·53 | 3·70 |
| Epistasis | C | 1·45 | 1·83 | 2·46 | 2·83 |
| | R | 1·23 | 1·57 | 2·05 | 2·63 |
| $G \times E$ (0·1) | C | 1·11 | 1·28 | 1·69 | 1·88 |
| | R | 1·04 | 1·32 | 1·44 | 1·96 |
| $G \times E$ (0·5) | C | 1·07 | 1·31 | 1·65 | 2·03 |
| | R | 0·89 | 0·93 | 1·31 | 1·49 |
| $G \times E$ (0·8) | C | 1·23 | 1·80 | 2·18 | 2·74 |
| | R | 1·11 | 1·51 | 1·66 | 2·06 |

(Population size refers to the number of individuals of each sex; C and R indicate total coupling and total repulsion initial gametic phase)

chromosome (from 6 to 51) or the number of markers in the index (from 3 to 15) has relatively little influence on the response to MAS. Even with only three markers per chromosome, the efficiency of MAS for characters with dominance or epistasis is surprisingly high, if six markers (out of a total of 30) are included in the index, although it is greatly reduced if only three markers are in the index.

Besides simulations in which markers included in the selection index were chosen in each generation out of all the markers in the genome, we have also conducted simulations in which markers included in the index were chosen in the first generation, and only their effects were reestimated in subsequent generations. The efficiency of MAS in such simulations was always much lower than in the simulations with all markers in the genome reevaluated each generation.

In all our simulations of MAS for characters with non-additive gene action as well as with non-additive environment, the coefficient $b_M$ in the selection index (6), i.e., the relative weight in the index of markers was so high that selection was based exclusively on the markers with the phenotype playing no role at all. The same was also true in simulations of MAS for purely additive characters (Gimelfarb & Lande, 1994). Actually for population sizes of at least 100 individuals of each sex the values of $p_M$ in (8) was always greater than one, making the denominator in (7) negative, and the computer program automatically excluded the phenotype from selection index. Even when $R^2$ in (8) was corrected as described in our previous paper, the value of $p_M$ still exceeded one. Using the narrow instead of the broad heritability in computing $b_M$ would not help. Indeed, in simulations with non-additive environment, both heritabilities were the same. In simulations with non-additive gene action, on the other hand, given that the narrow heritability is always lower than broad heritability, replacing the latter with the former in (8) will yield even higher values of $p_M$. Can it be that giving some weight to the phenotype in the selection index may improve the efficiency of MAS? We have conducted simulations in which the coefficient $b_M$ in the index was not computed based on formulas (7) and (8), but was fixed on a particular value: $b_M = 9$, 4 and 1 corresponding to 10%, 20% and 50% of the contribution by the phenotype in the index. In all simulations with the phenotype contributing 10%, the efficiency of MAS was virtually the same as in simulations in which phenotype played no role in selection, whereas in all simulations with contribution by the phenotype greater than 10%, the efficiency was lower. Thus, as for a purely additive character, in a population of at least 100 individuals of each sex MAS for characters with non-additive gene action or environment is most effective if it is based exclusively on markers.

In all our simulations, the method of Lande & Thompson (1990) was applied without any adjustments for non-additivity of the traits. It is possible that the linear model based strictly on the additive effects associated with markers is not adequate for traits that are non-additive. This seems to be the case particularly for characters with non-additive environment. Some adjustments for non-additivity can, in principle, be made. For example, not only the additive marker effects but also dominance effects as well as the effects of pairwise interactions between markers can be estimated by linear regression (Haley & Knott, 1992), and included in the computation of the selection index. It is unlikely, however, that this will improve the efficient of MAS in the case of non-additive environments. Even in the cases of dominance and epistasis, it is not certain that accounting for non-additive effects will necessarily help. Indeed, this would require estimating a very large number of parameters with an inevitable decrease of their statistical significance. Also, pairwise interactions may not be sufficient to account for epistasis involving multiple QTLs (e.g., the 'multiplicative' epistasis considered in this paper), and including in the regression of higher order interactions between markers may be necessary, which would further decrease the statistical significance of estimated parameters.

## 4. Conclusions

The main conclusion of this work is that MAS based on the 'additive effects' of genetic markers estimated by linear regression of the phenotype on marker genotypes can be efficient even if the action by genes or environment on the character is not additive. MAS for such characters is more effective than selection based strictly on the phenotype. The efficiency of MAS for characters with negative dominance is even higher than the efficiency of MAS for characters with purely additive genetic and environmental effects.

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

MAS for characters with non-additive environment is the least efficient.

Practically all conclusions reached in our previous work for additive characters are also valid for non-additive characters. Increasing the number of markers on a chromosome or the number of markers included in the computation of the selection index does not necessarily result in more effective selection. If markers included in the selection index are chosen in each generation out of the total number of markers in the genome, the efficiency of selection is much higher than if the markers in the index are chosen in the first generation with only their effects reestimated in the subsequent generations. For population sizes of at least 100 individuals of each sex, the efficiency of MAS is maximized if the phenotype does not contribute to the selection index, i.e., selection is based exclusively on genetic markers. Population size is the most important factor determining the efficiency of MAS.

This work was supported by U.S. Public Health Service grant GM27120.

## References

Haley, C. S. & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315–324.

Gimelfarb, A. (1989). Genotypic variation maintained under stabilizing selection without mutations: epistasis. *Genetics* **123**, 217–227.

Gimelfarb, A. & Lande, R. (1994). Simulation of marker assisted selection in hybrid populations. *Genetical Research* **63**, 39–47.

Lande, R. & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743–756.

Zhang, W. & Smith, C. (1992). Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theoretical and Applied Genetics* **83**, 813–820.

Zhang, W. & Smith, C. (1993). Simulation of marker-assisted selection utilizing linkage disequilibrium: the effects of several additional factors. *Theoretical and Applied Genetics* **86**, 492–496.

## Appendix

The initial variance components in simulations with non-additive gene effects (Table 2) were computed by the standard formulas utilizing the broad, $H^2$, and narrow, $h^2$, heritabilities in the initial population:

$$V_A/V_P = h^2, \quad \text{(A 1)}$$

$$V_E/V_P = 1-H^2, \quad \text{(A 2)}$$

$$V_{NA}/V_P = H^2-h^2, \quad \text{(A 3)}$$

where $V_P$ is the phenotypic variance in the initial population. The initial $H^2$ was 0·1 in all runs. The initial narrow heritability was estimated by the realized heritability in the first generation of purely phenotypic selection:

$$h^2 = R/S, \quad \text{(A 4)}$$

where $S$ and $R$ are, respectively, the average selection differential and the average response in the first generation of purely phenotypic selection for a number of runs with each particular set of parameter values (see text). The selection differential was computed by using the formula:

$$S = i\sigma_P, \quad \text{(A 5)}$$

where $i$ is the intensity of selection (standardized selection differential) and $\sigma_P = \sqrt{V_P}$ (Falconer 1983). Given the proportion of selected individuals 0·25 and assuming a normal distribution of the character in the initial population, $i = 1{\cdot}25$ (Falconer 1983, fig. 11.3). The response, $R$, was calculated directly as the difference between the mean in the first generation of phenotypic selection and the mean in the initial population.

As for the components of variance in the initial population with non-additive environment (Table 2), given the assumption of additive gene action, the additive genetic variance is the same as the genotypic variance, and, hence, $V_A = V_G$, $h^2 = H^2$, and $V_D = V_I = 0$. The genotypic and environmental variances are the variances of the components $G$ and $E$ in the decomposition of the phenotype,

$$Z = G+E+G\times E, \quad \text{(A 6)}$$

where $G$ is the genotypic value, $E$ is the environmental effect, and their sum, $G+E$, provides the least squares fit to the phenotype. Given that the phenotype is determined by expression (4), these components are, by definition, functions $G(z)$ and $E(e)$ producing the minimum of the integral

$$\int_z\int_e [(1-B)(z+e)+Bze-G(z) - E(e)]^2 p(z)\,q(e)\,dz\,de, \quad \text{(A 7)}$$

where $p(z)$ and $q(e)$ are the distributions of the genotypic value and of the environmental effect, respectively. It can be shown that the minimum of the integral is produced by

$$G = (1-B)z, \quad \text{(A 8)}$$

$$E = (1-B)e+Bem, \quad \text{(A 9)}$$

where $m$ is the mean genotypic value. Consequently,

$$G\times E = BE(z-m). \quad \text{(A 10)}$$

Since the mean genotypic value in the initial population is zero,

$$E = (1-B)e. \quad \text{(A 9}a\text{)}$$

$$G\times E = Bez. \quad \text{(A 10}a\text{)}$$

Given that $V_E+V_{GE} = (1-H^2)V_P$,

$$V_E/V_P = (1-H^2)/(1+\beta^2 v_z), \quad \text{(A 11)}$$

$$V_{GE}/V_P = (1-H^2)\beta^2 v_z/(1+\beta^2 v_z), \quad \text{(A 12)}$$

where $\beta = B/(1-B)$ and $v_z = V_G/(1-B)^2 = V_P H^2/(1-B)^2$ is the variance of the genotypic values in the initial population.

It follows from (A 11) that for a given broad

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press

heritability among $F_2$ offspring, the variance of the environmental effect is

$$v_e = (1 - H^2)\, V_P/[(1 - B)^2 + B^2 v_z], \qquad \text{(A 13)}$$

where $V_P$ and $v_z$ are, respectively, the phenotypic variance and the variance of genotypic values expected among $F_2$ offspring. The latter variances were estimated among 10000 offspring generated from the initial population at the beginning of each computer simulation.

https://doi.org/10.1017/S0016672300032730 Published online by Cambridge University Press