Marker-Assisted Selection as a Strategy for Increasing the Probability of Selecting Superior Genotypes

Steven J. Knapp

ABSTRACT

Marker-assisted selection (MAS) has been shown, in theory, to produce greater selection gains than phenotypic selection for normally distributed quantitative traits. Theory is presented in this paper for estimating the probability of selecting one or more superior genotypes by MAS (Pr_{MAS}). This paramater was used to estimate the cost efficiency of MAS relative to phenotypic selection (E_c). Pr_{MAS} and E_c are functions of heritability (h^2), heritability of a MAS index (h_I^2), the phenotypic selection threshold (i), the genotypic superiority threshold (g), and $p = \sigma_M^2/\sigma_G^2$, where σ_M^2 is additive genetic variance associated with markers and σ_G^2 is additive genetic variance. h_I^2 increases as p increases. Heritability can be increased to 1.0 by increasing p to 1.0; however, estimated marker effects ($\hat{p}$) and true quantitative trait locus effects (p) must be perfectly correlated to achieve this in practice. Pr_{MAS} increases throughout the range of p when $i \geq g$, decreases as g increases, and increases as i increases for most p. The frequency of superior genotypes among selected progeny increases as selection intensity increases. E_c ranged from 1.0 to 16.7 for i and g from 1.282 to 2.326, h^2 from 0.1 to 1.0, and p from 0.0 to 1.0; thus, a breeder using phenotypic selection must test 1.0 to 16.7 times more progeny than a breeder using MAS to be assured of selecting one or more superior genotypes. E_c increases as g or i increase and h^2 decreases and increases as p increases when $i = g$. E_c predicts that MAS substantially decreases the resources needed to accomplish a selection goal for a low to moderate heritability trait when the selection goal and the selection intensity are high.

THE PROBABILITY of selecting superior genotypes is low for low to moderate heritability (h^2) traits (Robson et al., 1967; Johnson, 1989). Plant breeders cope with this problem by producing and testing progeny from numerous crosses, using low selection intensities, using replicated testing, testing advanced generations, and using recurrent selection (either truly cyclic population improvement schemes or "second generation" crosses between "suboptimum" inbred lines) (Hallauer and Miranda, 1981). Marker-assisted selection (MAS) has emerged as a strategy for increasing selection gains (Dudley, 1993; Lande and Thompson, 1990; Lande, 1992; Knapp, 1994a). Although the gains from marker-assisted index selection are theoretically greater than the gains from phenotypic selection (Lande and Thompson, 1990), quantitative trait locus (QTL) and MAS index parameter estimation errors, genetic drift, and disequilibrium between selected and unselected QTL can reduce the gains from MAS and may lead to lower selection gains for MAS than for phenotypic selection, particularly in long range or recurrent selection experiments (Beavis, 1994, 1997; Bulmer, 1971; Dudley, 1993; Gimelfarb and Lande, 1994a,b, 1995; Knapp et al., 1993; Knapp, 1994b; Lande and Thompson, 1990; Lande, 1992; Zhang and Smith, 1992, 1993).

MAS should be most effective in the early generations of selection among progeny from crosses between inbred lines (Lande, 1992; Stromberg et al., 1994). Heritabilities are usually lowest (because replications are limited and experimental units tend to be small) and linkage disequilibrium is greatest in these generations (Falconer, 1981). The paradox is that the power for mapping QTL decreases as heritability decreases and is lowest for traits where MAS has the greatest theoretical impact (Lande and Thompson, 1990; Lande, 1992). The accuracy of QTL and MAS index parameter estimates can be low when heritability is low and samples are small (Beavis, 1994, 1997; Gimelfarb and Lande, 1995). This problem is not unique to early generation MAS. Early generation phenotypic selection is seldom strongly advocated in crop plants despite the theoretical drawbacks of delaying selection (Geiger, 1984; Snape and Simpson, 1984; Sneep, 1977, 1984; Weber, 1984). Selection is frequently delayed to later generations because heritabilities and the statistical accuracy of progeny mean estimates tend to increase as the number of replications, generations, sites, and years of testing increase.

Selecting in the early generations of a pedigree breeding program poses special problems (Geiger, 1984; Snape and Simpson, 1984). Seed supplies are often limited and the chance of advancing superior genotypes through the early generations (F_2 and F_3) is low for some traits. Mean performance across sites and years is often poorly estimated from F_2 and F_3 phenotypic observations, and these observations may be from less than optimum experimental units (hills or small plots). Limited resources, however, dictate either discarding a large fraction of lines early or delaying selection until the F_4 or later and testing fewer progeny in the process. The second strategy forces breeders to distribute resources to a smaller number of progeny tested across a larger number of generations, sites, and years (Geiger, 1984; Snape and Simpson, 1984). The trade-off is between producing more accurate estimates of progeny means versus sampling a larger number of progeny per cross.

Although organisms, traits, and circumstances differ greatly, there are two universal sampling problems in breeding programs. First, enough progeny must be tested and selected to ensure that at least one has a superior genotype (is fixed for more favorable alleles than the parents or has a genotypic mean exceeding

Crop Sci. Bldg. 451C, Dep. of Crop and Soil Science, Oregon State Univ., Corvallis, OR 97331. This work was funded by a grant from the USDA NRICGP (90-43213-5190). Oregon Agric. Exp. Stat. Tech. Paper No. 11,198. Received 30 June 1997. *Corresponding author (knapps@css.orst.edu).

Published in Crop Sci. 38:1164–1174 (1998).

Abbreviations: MAS, marker-assisted selection; QTL, quantitative trait locus.

a genotypic superiority threshold selected by the breeder). When the heritabilities of the selected traits are low or moderate and small samples of progeny are tested, the probability of selecting an outstanding genotype is very low (Robson et al., 1967; Johnson, 1989). Second, selected progeny are mixtures of inferior and superior genotypes. The frequency of inferior genotypes in a selected sample of progeny increases as heritability decreases (Robson et al., 1967). The usual strategy for sorting superior from inferior genotypes is ''advanced testing.'' The problem with advanced testing strategies is that the most outstanding genotypes are often not selected in early generations when heritabilities are low or samples are small or both and thus are not present among progeny selected for advanced testing.

Robson et al. (1967) described methods for estimating the frequency of superior genotypes in samples of progeny (superior phenotypes) selected for normally distributed quantitative traits using phenotypic selection. This study highlighted some of the dilemmas faced by breeders selecting for low to moderate heritability traits. First, very large samples are often needed to ensure the presence of one or more superior genotypes in the selected sample. Plant breeders seldom test enough progeny from one cross to be assured of retaining outstanding genotypes when heritabilities are low (Johnson, 1989). Second, large numbers of progeny must be selected (low selection intensities must be used) to ensure the presence of one or more superior genotypes in the selected sample. Even when low selection intensities are used, the most outstanding genotypes produced by a cross might not be present in the selected sample when heritability is low and samples are small (Robson et al., 1967; Johnson, 1989). With unlimited resources, breeders would (i) delay selection until sufficient testing had been done to identify, with some degree of accuracy, the most outstanding genotypes from each cross and (ii) test a sufficient number of progeny to ensure the presence of one or more superior genotypes among the selected progeny.

If breeders had tools to increase heritability cost effectively, then breeding program throughputs and efficiency could be greatly increased by testing fewer progeny per cross, culling inferior progeny early, and using higher selection intensities. The problem with implementing MAS, apart from QTL parameter estimation errors, is the cost difference between molecular marker and phenotypic assays for most traits. This difference should steadily decrease as the technology advances (Perlin et al., 1995; Schwengel et al., 1994; Vos et al., 1996), and advances in the technology should increase the merit of MAS as a strategy for increasing heritability. Theory is presented in this paper for calculating (i) the heritability of a MAS index (h_I^2), (ii) the probability of selecting one or more superior genotypes by MAS (Pr_{MAS}), and (iii) the efficiency of MAS relative to phenotypic selection (E_c) for normally distributed quantitative traits. One of my primary aims was to compare the number of progeny a breeder needs to test to be assured of selecting one or more superior genotypes when using phenotypic or marker assisted index selection. By comparing these numbers and the cost per experimental observation, breeders can calculate the difference in cost between phenotypic and marker assisted index selection. Additionally, these numbers can be used as the basis for planning QTL mapping and selection experiments.

THEORY

Heritability of the MAS Index

Lande and Thompson (1990) described an optimum index for selecting individuals or lines (families) for a normally distributed quantitative trait. This index is a weighted sum of phenotypic and marker scores, with weights calculated as per an optimum selection index (Hazel, 1943; Smith, 1936). The vector of index scores for one trait is estimated by $\mathrm{I} = \mathrm{b}_p\mathrm{x} + \mathrm{b}_M\mathrm{m}$, where

$$\mathrm{b} = P^{-1}Gd = \begin{bmatrix} b_P \\ b_M \end{bmatrix} \qquad [1]$$

is a vector of index weights, x is an $N \times 1$ vector of phenotypic scores, $\mathrm{m} = \Sigma_k \hat{\alpha}_k n_k$ is an $N \times 1$ vector of marker scores, N is the number of progeny tested, $\hat{\alpha}_k$ is the additive effect of the kth marker locus, n_k is the number of favorable alleles at the kth marker locus,

$$b_P = \frac{\sigma_G^2 - \sigma_M^2}{\sigma_P^2 - \sigma_M^2} = \frac{1 - p}{1/h^2 - p}$$

is the index coefficient for phenotypic scores,

$$b_M = \frac{\sigma_P^2 - \sigma_G^2}{\sigma_P^2 - \sigma_M^2} = \frac{1/h^2 - 1}{1/h^2 - p}$$

is the index coefficient for marker scores, σ_G^2 is the additive genetic variance between individuals or lines, σ_P^2 is the phenotypic variance between individuals or lines, σ_M^2 is the additive genetic variance associated with marker loci, $p = \sigma_M^2/\sigma_G^2$ is the proportion of the additive genetic variance associated with markers, $h^2 = \sigma_G^2/\sigma_P^2$ is the heritability,

$$P = \begin{bmatrix} \sigma_P^2 & \sigma_M^2 \\ \sigma_M^2 & \sigma_M^2 \end{bmatrix}$$

is the phenotypic variance-covariance matrix,

$$G = \begin{bmatrix} \sigma_G^2 & \sigma_M^2 \\ \sigma_M^2 & \sigma_M^2 \end{bmatrix}$$

is the genotypic variance-covariance matrix, and

$$\mathrm{d} = \begin{bmatrix} d_P \\ d_M \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

is the economic weight vector (Lande and Thompson, 1990).

The heritability of the optimum index for two or more normally distributed quantitative traits is

$$h_I^2 = \frac{b'G\mathrm{b}}{b'P\mathrm{b}}$$

(Lin and Allaire, 1976). The heritability of the MAS index for a normally distributed quantitative trait is

$$h_I^2 = \frac{b'G\mathrm{b}}{b'P\mathrm{b}} = \frac{\sigma_G^2}{\sigma_P^2 - \sigma_M^2} + \frac{\sigma_M^2}{\sigma_M^2 - \sigma_P^2} + \frac{\sigma_M^2(\sigma_P^2 - \sigma_G^2)}{\sigma_G^4 - 2\sigma_G^2\sigma_M^2 + \sigma_M^2\sigma_P^2}$$
$$= \frac{1}{1/h^2 - p} + \frac{p}{p - 1/h^2} + \frac{p(1 - h^2)}{h^2 - 2ph^2 + p}, \qquad [2]$$

where b, P, and G are from Eq. [1]. The calculations shown in this paper use Eq. [2] and apply to virtually any selection scheme (e.g., individual, family, or line selection) for a normally distributed quantitative trait. Although the experimental and genetic definitions of σ^2_G, σ^2_M, and σ^2_P change as breeding schemes (e.g., individual versus family selection and pedigree versus half-sib family selection) and experiment and environment designs (Falconer, 1981; Wricke and Weber, 1986) change, the calculations shown in this paper depend on h^2 and p (not on the magnitudes of σ^2_G, σ^2_M, and σ^2_P) and can be applied to any breeding scheme or mating design where selection on the MAS index can be used (Lande, 1992). All of the calculations directly compare the choice between selecting on phenotypic versus MAS index scores for samples of progeny produced by the same mating, experiment, and environment designs.

Segregating populations were simulated to illustrate principles underlying the theoretical calculations and the theoretical impact of MAS on genotypic effect (G_j) distributions. Five hundred progeny were simulated for factorial combinations of p = 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 by h^2 = 0.1, 0.2, and 0.5 for a normally distributed quantitative trait (one sample was simulated for each h^2 and p combination). Genotypic effects (Gj) and non-genotypic (Ej) effects were simulated with the RANNOR function of SAS (1985) and summed to produce phenotypic effects ($I_j = G_j + E_j$). The expected variances supplied to the RANNOR function were calculated for each p and h^2 by setting $\sigma^2_E = 1$ and finding σ^2_G with Eq. [2], where $\sigma^2_P = \sigma^2_G + \sigma^2_E$ and σ^2_E is the non-genetic variance. Random seed numbers were supplied to RANNOR for each sample.

The Probability of Selecting Superior Genotypes

The probability of selecting at least one progeny with a genotypic value (G_j) greater than g' among progeny with phenotypic values (X_j) greater than x' is

$$Pr_{PS} = Pr[(G_j > g')|(X_j > x')] = \frac{1}{1 - \Phi(x)} \int_x^\infty \Phi\left[\frac{hz - g}{\sqrt{1 - h^2}}\right] d\Phi(z), \quad [3]$$

where x' is the phenotypic selection threshold, g' is an unobserved genotypic superiority threshold, $x = (x' - \mu)/\sigma_P$, $g = (g' - \mu)/\sigma_G$, Φ is the standard normal cumulative distribution function, and $\Phi(x)$ is the area under a standard normal distribution below x (Robson et al., 1967). Robson et al. (1967) tabulated Pr_{PS} for several g', x', and h^2. Pr_{PS} reduces to $[1 + \Phi(y)]/2$ for $h^2 = 0.5$ and $g = 0$, but must be numerically integrated for other g and h^2.

The number of progeny a breeder needs to test to be $100(1 - c)$% certain of selecting at least one superior genotype using phenotypic selection is

$$n_{PS} = \frac{\log_{10}(c)}{\log_{10}[1 - Pr_{PS}(1 - \Phi[x])]}, \quad [4]$$

where $(1 - c)$ is the assurance probability. This probability sets the number of samples greater than or equal to n_{PS} that should produce one or more superior genotypes, e.g., 99 out of 100 samples greater than or equal to n_{PS} should have one or more superior genotypes when c = 0.01.

The probability of selecting at least one progeny with a genotypic value (G_i) greater than g' among progeny with MAS index values (I_j) greater than i' is

$$Pr_{MAS} = Pr[(G_j > g')|(I_j > i')] = \frac{1}{1 - \Phi(i)} \int_i^\infty \Phi\left[\frac{h_I z - g}{\sqrt{1 - h_I^2}}\right] d\Phi(z), \quad [5]$$

where $i = (i' - \mu)/\sigma_I$, $\Phi(i)$ is the area under a standard normal distribution below i, and i' is the MAS index selection threshold. This probability was found by substituting Eq. [2] for h^2 in Eq. [3]. The number of progeny a breeder needs to test to be $100(1 - c)$% certain of selecting at least one superior genotype using MAS is

$$n_{MAS} = \frac{\log_{10}(c)}{\log_{10}[1 - Pr_{MAS}(1 - \Phi[i])]}. \quad [6]$$

The efficiency of MAS relative to phenotypic selection can be estimated by

$$E_c = n_{PS}/n_{MAS} = \frac{\log_{10}[1 - Pr_{MAS}(1 - \Phi[i])]}{\log_{10}[1 - Pr_{PS}(1 - \Phi[x])]}. \quad [7]$$

E_c can be used to assess whether or not MAS is a cost efficient for a specific breeding problem by comparing the cost per observation for phenotypic (c_{PS}) and marker (c_{MAS}) assays along with n_{PS} and n_{MAS}, e.g., if the cost per observation is ten times greater for MAS than for phenotypic selection (c_{MAS}/c_{PS} = 10) and E_c = 5, then phenotypic selection is twice as cost efficient as MAS ($(c_{MAS}/c_{PS})/(n_{PS}/n_{MAS})$ = 10/5 = 2) even though phenotypic selection requires five times as many progeny as MAS (E_c = 5) to achieve the same breeding goal. Although n_{PS} and n_{MAS} depend on the assurance probability $(1 - c)$, E_c is not affected by this variable and can be calculated by varying h^2, p, x, i, and g alone.

The calculations shown in this paper were done with the INTEGRATE function of Mathematica (Wolfram, 1989). Pr_{PS} and Pr_{MAS}, n_{PS} and n_{MAS}, and E_c were calculated for factorial combinations of three genotypic superiority thresholds (g = 1.282, 1.645, and 2.326), three phenotypic selection thresholds (x = 1.282, 1.645, and 2.326) or index selection thresholds (i = 1.282, 1.645, and 2.326), p between 0.0 and 1.0, and h^2 between 0.1 and 1.0. The thresholds 1.282, 1.645, and 2.326 are truncation points for the upper 10, 5, or 1%, respectively, of the phenotypic, index, and genotypic distributions. The same selection intensities were used for MAS and phenotypic selection ($x = i$). n_{PS} and n_{MAS} were calculated with assurance probabilities $(1 - c)$ of 0.80, 0.90, 0.95, 0.99 and 0.999.

RESULTS

Heritability of the MAS Index

The heritability of the MAS index (h^2_I) was calculated for p from 0.0 to 1.0 for 10 initial heritabilities (h^2)

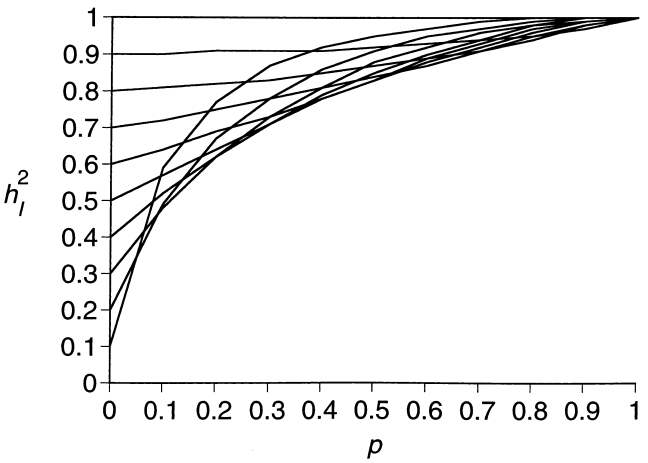

Fig. 1. Heritability of the MAS index (h^2_I) for $p = \sigma^2_M/\sigma^2_G$ ranging from 0.0 to 1.0 and heritability (h^2) ranging from 0.1 to 1.0, where σ^2_M is the additive genetic variance associated with markers and σ^2_G is the additive genetic variance.

ranging from 0.1 to 1.0 (Fig. 1). When MAS is not used, $p = 0$ and the heritability of the index reduces to the heritability for phenotypic selection ($h_I^2 = h^2 = \sigma_G^2/\sigma_P^2$) (Fig. 1). h_I^2 increases as p increases when $h^2 < 1$ and can, in theory, be increased to 1.0 by increasing p to 1.0. Most of the heritability increase produced by the MAS index accrues between $0.0 < p < 0.5$ for low to moderate heritability traits ($0.0 < h^2 < 0.7$) (Fig. 1). The effect of p on h_I^2 is non-linear and increases as h^2 decreases (Fig. 1). The non-linearity and h_I^2 by p cross-over interaction is caused by the effects of the index weights (Fig. 1). The ranges of the index weights are $h^2 \leq b_P \leq 0$ and $(1 - h^2) \leq b_M \leq 1$ for $0 \leq p \leq 1$, where $b_P + b_M = 1$ (Lande and Thompson, 1990); thus, b_P decreases, whereas b_M increases as h^2 decreases. When the heritability of one trait is greater than the heritability of another trait and p is the same for both traits, h_I^2 is greater for the trait with the lower heritability (Fig. 1).

The Probability of Selecting Superior Genotypes

Pr_{MAS} plots are shown for factorial combinations of g and i for three heritabilities (Fig. 2). Pr_{MAS} increases to 1.0 as p increases for every h^2, i, and g and rapidly reaches 1.0 for some p, h^2, i, and g (Fig. 2). Pr_{MAS} increases as selection intensity increases (i increases) for most p. The frequency of superior genotypes increases among the selected progeny as selection intensity increases. Pr_{MAS} decreases as the genotypic superiority threshold (g) increases. The probability of selecting a superior genotype is greater, for example, when the goal is to select a genotype from the upper 10% versus the upper 1% of the genotypic distribution.

Pr_{MAS} increases throughout the range of p when $i = g$ (Fig. 2). Pr_{MAS} increases to 1.0 as p increases to 1.0 when $i \geq g$, but plateaus below 1.0 when $i < g$. The maximum Pr_{MAS} is $[\Phi(g)]/[\Phi(i)]$. The percentage of superior genotypes in the selected sample never exceeds this maximum; thus, the phenotypic selection threshold (i) must be greater than or equal to the genotypic superiority threshold (g) for Pr_{MAS} to reach 1.0, e.g., when $g = 2.236$ and $i = 1.282$, the frequency of superior genotypes in the selected sample cannot exceed $[\Phi(g)]/[\Phi(i)] = 0.01/0.10 = 0.10$ (10%). Pr_{MAS} rapidly reaches this threshold as p increases when heritability is low and $i < g$ (Fig. 2).

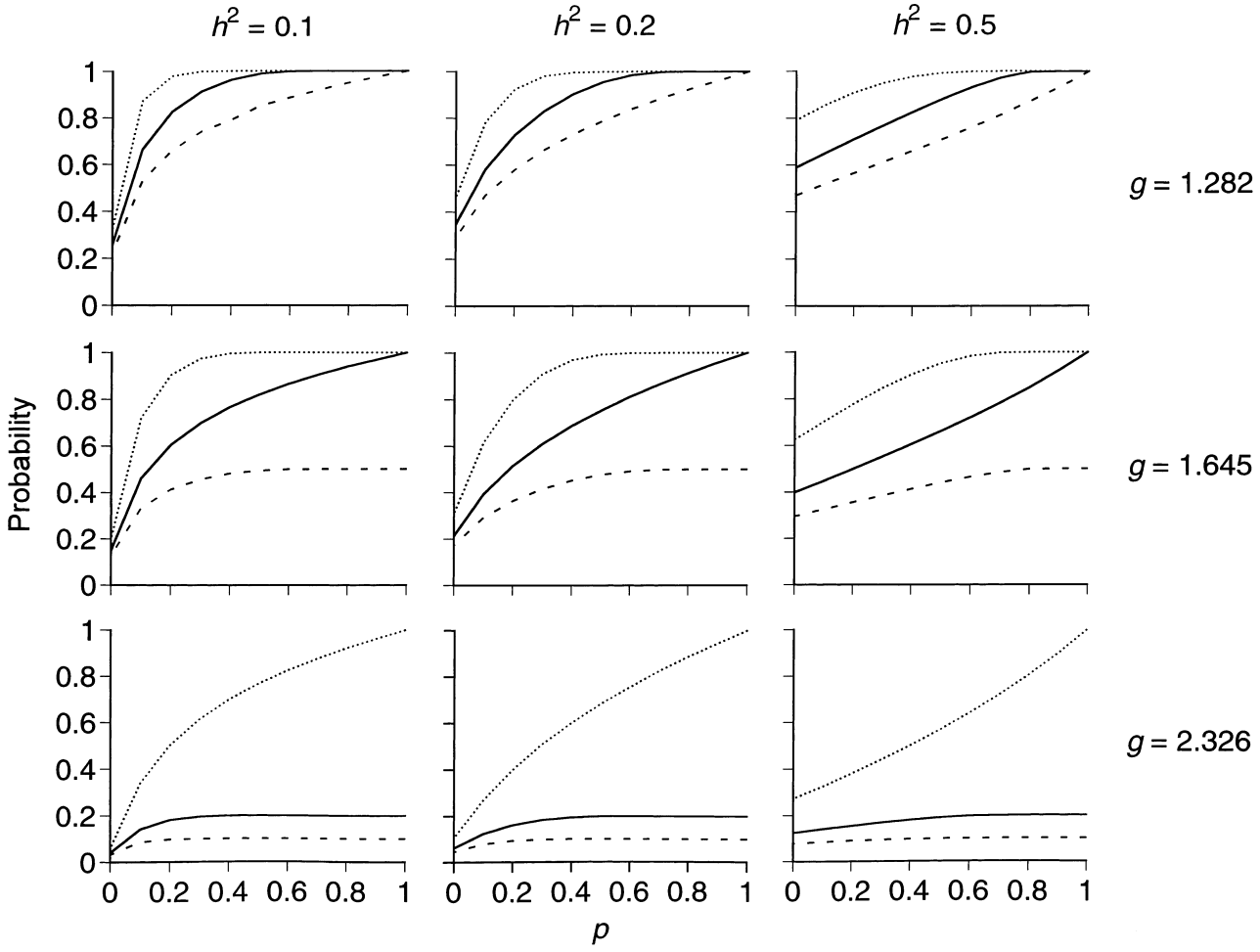

Fig. 2. The probability of selecting one or more superior genotypes using phenotypic selection ($p = 0$) or marker-assisted index selection ($p > 0$) for a normally distributed quantitative trait. Probabilities are shown for three MAS index selection thresholds, $i = 1.282$ (dotted lines), $i = 1.645$ (solid lines), and $i = 2.326$ (dashed lines), three genotypic superiority thresholds, $g = 1.282$ (upper row), $g = 1.645$ (middle row), and $g = 2.326$ (lower row), and three heritabilities, $h^2 = 0.1$ (left column), $h^2 = 0.2$ (middle column), and $h^2 = 0.5$ (right column), where $p = \sigma_M^2/\sigma_G^2$,$\sigma_M^2$ is the additive genetic variance associated with markers, σ_G^2 is the additive genetic variance, g is the genotypic standard deviation, and i is the MAS index standard deviation.

The frequency of superior genotypes among the selected progeny is greater for MAS than for phenotypic selection for most h^2, p, g, and i (Fig. 2). The differences between phenotypic selection and MAS are most dramatic for low heritability traits, high selection intensities, and high genotypic superiority thresholds. The differences can still be dramatic for moderate heritabilities and modest selection goals.

The Efficiency of MAS Relative to Phenotypic Selection

Efficiency plots are shown for factorial combinations of g and i for three heritabilities (Fig. 3). E_c ranged from 1.0 to 16.7 for i from 1.282 to 2.326, g from 1.282 to 2.326, h^2 from 0.1 to 1.0, and p from 0.0 to 1.0. Efficiency increases as h^2 decreases independent of i and g and as p increases when $i = g$ (Fig. 3). The efficiency maximums for $i \neq g$ are less than for $i = g$. Setting the selection intensity lower than the genotypic selection threshold ($i < g$) is less efficient than setting $i = g$ because the frequency of inferior genotypes in the selected fraction increases as i decreases (Fig. 3). Conversely, setting the selection intensity higher than the genotypic selection threshold ($i > g$) is less efficient than setting $i = g$ for some p because i affects the speed with which Pr_{MAS} reaches 1.0. E_c plateaus once Pr_{MAS} reaches 1.0 (Fig. 3).

The effect of selection intensity on E_c decreases as the genotypic superiority threshold (g) decreases for most h^2 (Fig. 3). MAS is most efficient when breeders use high selection intensities (e.g., select progeny from the upper 1% of the population) and set high selection goals (e.g., require that the experiment produce at least one selection from the upper 1% of the genotypic distribution). Selection intensity must be increased to exclude inferior genotypes when heritability is increased by using MAS. Low selection intensities must be used when heritabilities are low because high selection intensities will frequently exclude the most outstanding genotypes.

Some of the principles underlying the efficiency calculations are illustrated in a series of plots showing the genotypic and phenotypic ($p = 0$) or MAS index ($p > 0$) effect distributions for simulated segregating populations for low to moderate heritability traits ($h^2 = 0.1$, 0.2, and 0.5) (Fig. 4–6). The most outstanding genotypes (those with the highest genotypic values) in the samples for the 10 or 20% heritability traits would not be selected by phenotypic selection alone ($p = 0$), whereas the most outstanding genotypes in all of the MAS samples ($p > 0$) would almost certainly be selected. The phenotypic ranks of the top three genotypes in the phenotypic selection samples were 479, 57, and 67 for $h^2 = 0.1$ (Fig. 4), 241, 11, and 88 for $h^2 = 0.2$ (Fig. 5), and 87, 3, and 28 for $h^2 = 0.5$ (Fig. 6). The top ranking genotype would not be selected from the sample for the 10% heritability trait without retesting the whole population, would only be selected from the sample for the 20% heritability trait by keeping nearly 50% of the progeny for addi-

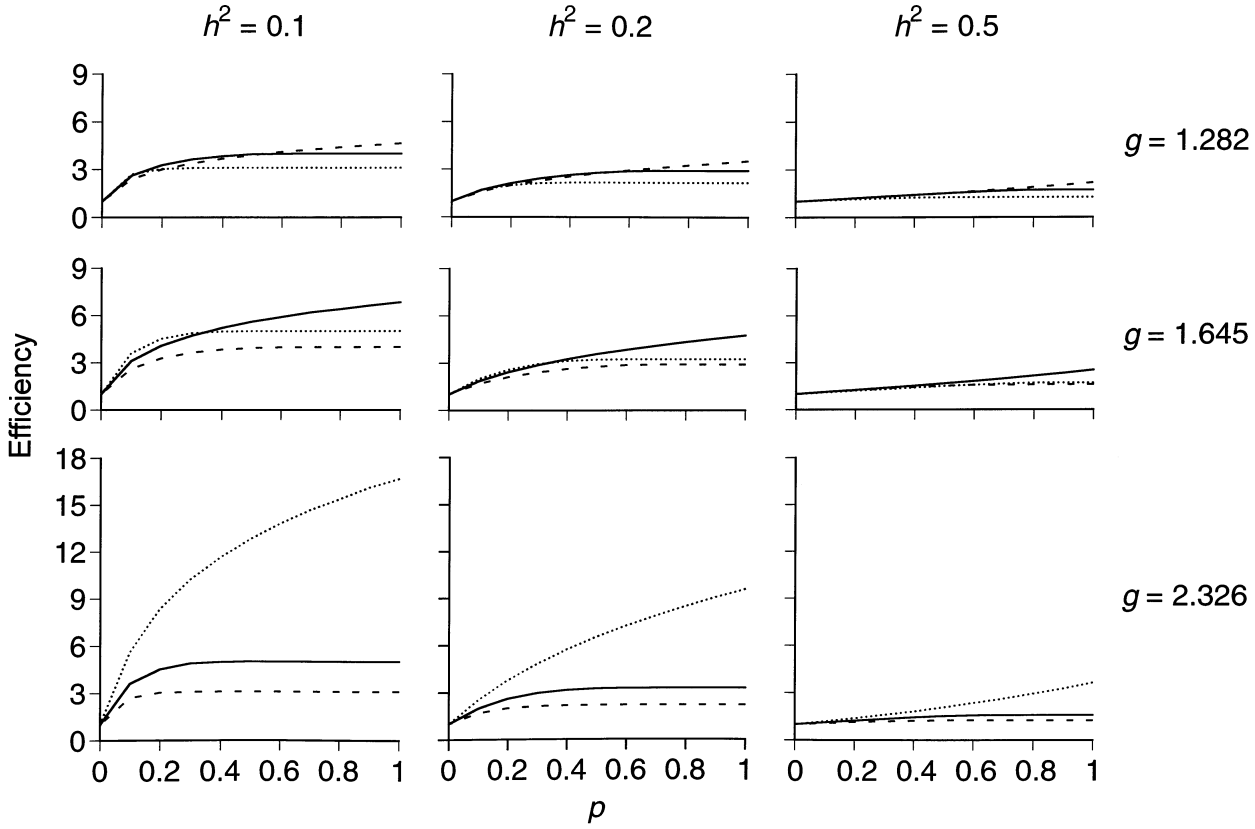

Fig. 3. The efficiency ($E_c = n_{PS}/n_{MAS}$) of marker-assisted index selection relative to phenotypic selection for a normally distributed quantitative trait for three MAS index selection thresholds, i = 1.282 (dotted lines), i = 1.645 (solid lines), and i = 2.326 (dashed lines), three genotypic superiority thresholds, g = 1.282 (upper row), g = 1.645 (middle row), and g = 2.326 (lower row), and three heritabilities, h^2 = 0.1 (left column), h^2 = 0.2 (middle column), and h^2 = 0.5 (right column), where $p = \sigma^2_M/\sigma^2_G$, σ^2_M, is the additive genetic variance associated with markers, σ^2_G is the additive genetic variance, g is the genotypic standard deviation, and i is the MAS index standard deviation.

tional testing, and would only be selected from the sample for the 50% heritability trait by keeping 17% of the progeny for additional testing. The second ranking genotypes from the samples for the 20 and 50% heritability traits, however, would be selected with high selection intensities (2.0 and 0.4%, respectively).

Suppose the upper 10% of the progeny (50 of 500) are selected by phenotypic selection ($p = 0$). Fourteen, eight, and one progeny are from the upper 10, 5, and 1%, respectively, of the genotypic distribution for $h^2 = 0.1$ (Fig. 4), 12, eight, and two progeny are from the upper 10, 5, and 1%, respectively, of the genotypic distribution for $h^2 = 0.2$ (Fig. 5), and 20, 11, and three progeny are from the upper 10, 5, and 1%, respectively, of the genotypic distribution for $h^2 = 0.5$ (Fig. 6); thus, most of the selected progeny (30 to 49 of 50) are inferior when progeny are selected by phenotypic scores alone and h^2 ranges from 0.1 to 0.5. These examples illustrate two problems: the most outstanding (the top ranking) genotypes are not selected and a significant fraction of the progeny carried forward for additional testing are inferior when heritabilities are low, e.g., when the goal is to select a line from the upper 1% of the genotypic distribution and the upper 10% of the phenotypes are selected, 94 to 98% of the selected progeny are inferior when h^2 ranges from 0.1 to 0.5.

The outcome of selection can be enhanced by using a MAS index, even when only a modest proportion of the additive genetic variance is associated with markers (Fig. 4–6). This is illustrated by the dramatic changes in the index (phenotypic) distributions between $p = 0.0$ and $p = 0.2$ (Fig. 4–6). The observed heritabilities increased from $\hat{h}^2 = 0.112$ for $p = 0.0$ to $\hat{h}_I^2 = 0.769$ for $p = 0.2$ in the samples simulated for a 10% heritability trait (Fig. 4) and from $\hat{h}^2 = 0.184$ for $p = 0$ to $\hat{h}_I^2 = 0.691$ to for $p = 0.2$ in the samples simulated for a 20% heritability trait (Fig. 5). The genotypic ranks of the upper 1% (5 of 500) of the progeny selected with MAS index scores with $p = 0.20$ were one, 8, 7, 4, and 31 for the $h^2 = 0.1$ trait (Fig. 4), one, two, five, 11, and 21 for the $h^2 = 0.2$ trait (Fig. 5), and one, five, 11, 21, and 92 for the $h^2 = 0.5$ trait (Fig. 6). The top ranking genotypes are the top ranking phenotypes in these samples and would be selected with any selection intensity.

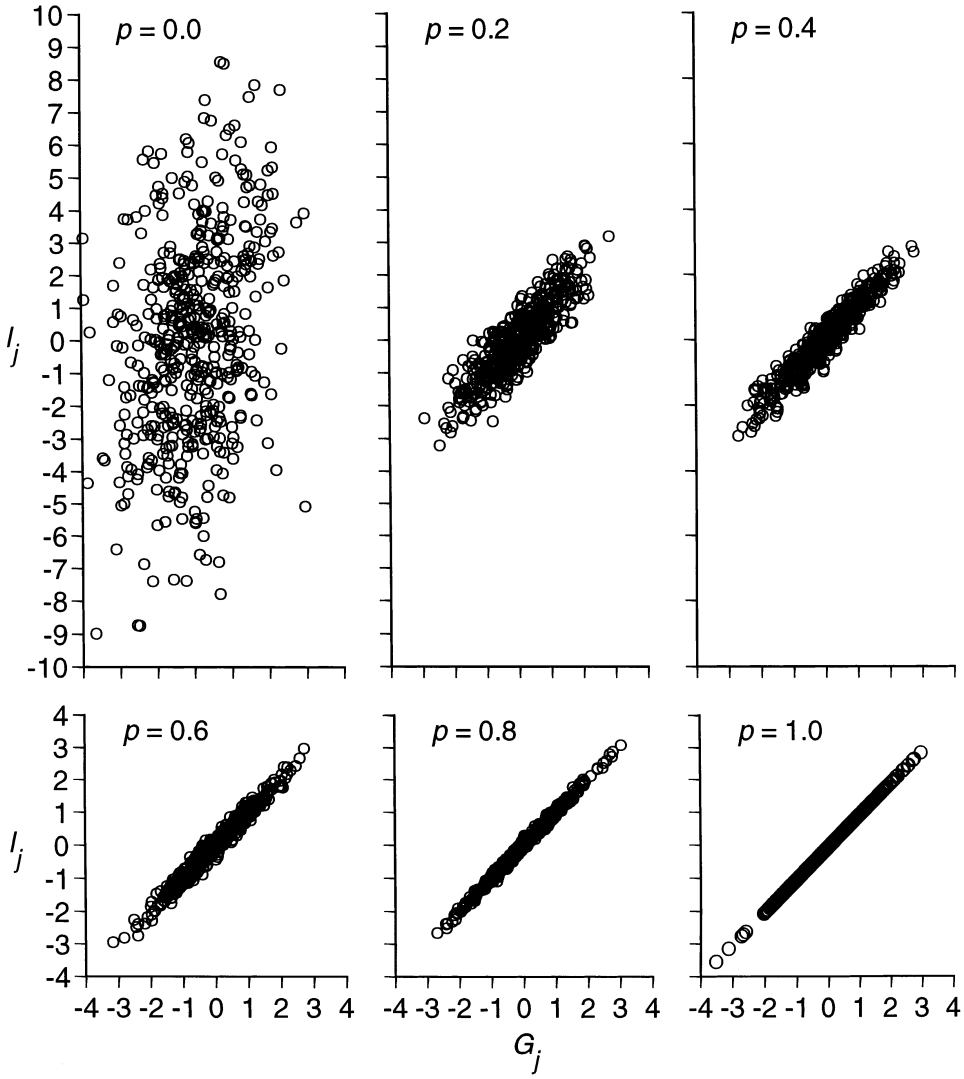

Fig. 4. Genotypic (G_j) and MAS index (I_j) effects for samples of 500 progeny simulated for a trait with $h^2 = 0.1$ for phenotypic selection ($p = 0$) and marker-assisted selection ($p = 0.2$ to 1.0 by 0.2). The estimated heritabilities ($\hat{h}^2$) for the samples simulated using $p = 0.0$, 0.2, 0.4, 0.6, 0.8, and 1.0 samples were 0.112, 0.769, 0.921, 0.967, 0.992, and 1.0, respectively.

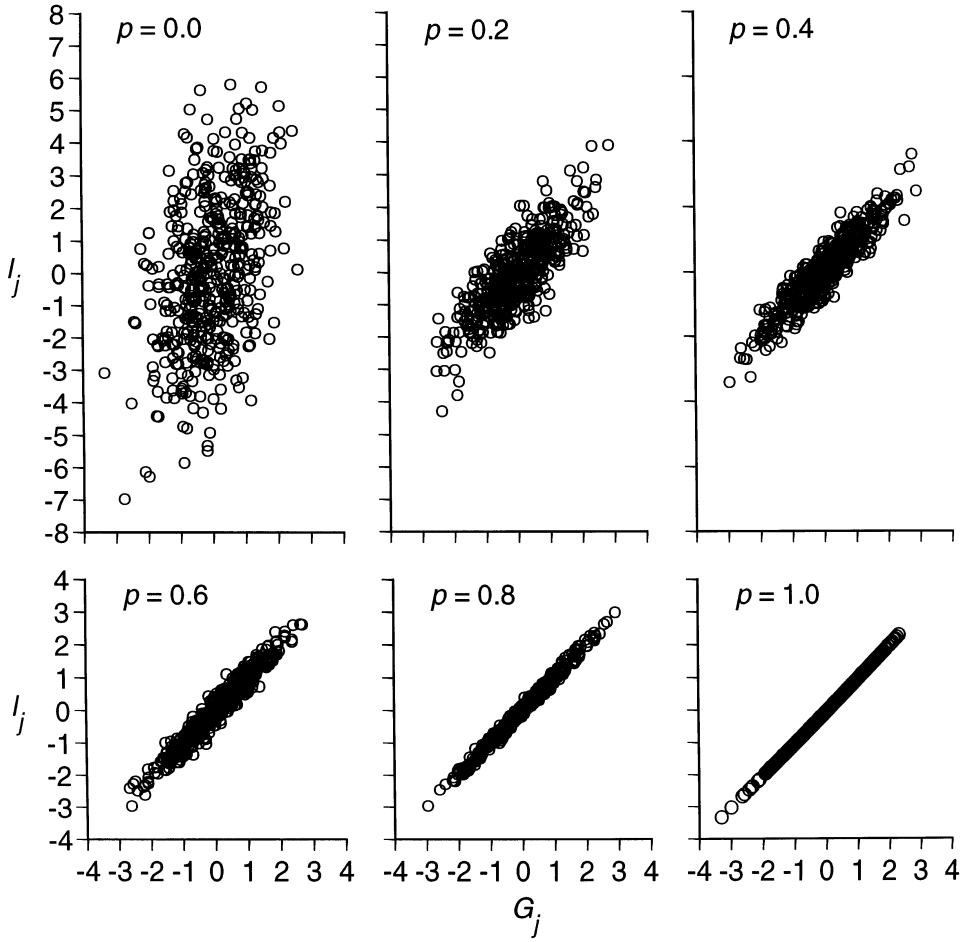

Fig. 5. Genotypic (G_j)) and MAS index (I_j) effects for samples of 500 progeny simulated for a trait with h^2 = 0.2 for phenotypic selection (p = 0) and marker-assisted selection (p = 0.2 to 1.0 by 0.2). The estimated heritabilities ($\hat{h}^2$) for the samples simulated using p = 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 samples were 0.184, 0.691, 0.875, 0.928, 0.992, and 1.0, respectively.

MAS Reduces the Resources Needed for Progeny Testing

n_{PS} and n_{MAS} are affected by the assurance probability $(1 - c)$. c must be chosen by the breeder (along with g and i) when calculating n_{PS} and n_{MAS}. n_{PS} and n_{MAS} increase as $(1 - c)$ increases (Table 1). When the heritability of the selected trait is 10%, the goal is to select one or more progeny from the upper 1% of the genotypic distribution (g = 1.282), and the upper 10% of the phenotypes (i = 1.282) are selected, twice as many progeny (203 versus 102) must be tested from each cross to be 99 as opposed to 90% certain of selecting one or more superior genotypes; thus, the choice of c affects how resources are allocated within and between crosses in a breeding program (Table 1). Although the means and variances of crosses vary, the overall frequency of superior genotypes in a breeding program is bound to be similar for different resource allocation strategies unless a preponderance of the crosses are inferior (have lower means and variances). Under such circumstances, breeding program resources would be disproportionately allocated to inferior progeny. Most breeders distribute risk across crosses because they lack a basis for confidently choosing between crosses (Dudley, 1984; St. Martin et al., 1996). If the merits of crosses are greatly different and there is a sound basis for choosing the most promising crosses, then the most efficient strategy is to concentrate resources on fewer crosses with more progeny per cross because this increases the probability of selecting superior genotypes (Table 1).

The effect of p on n_{MAS} was assessed by an assurance probability of 99% (Table 2). Substantially fewer progeny are needed for MAS (n_{MAS}) than for phenotypic selection (n_{PS}) to reach the same selection goal for most p, g, i, and h^2 (Table 2). n_{MAS} rapidly plateaus for many g, i, and h^2, changes most dramatically in the range $0.0 < p < 0.4$, and does not plateau for high i and g. MAS typically requires many fewer progeny than phenotypic selection to reach the same selection goal even when a small fraction of the additive genetic variance is associated with markers.

n_{PS} and n_{MAS} are profoundly affected by the goal of the breeder, e.g., when progeny are selected from the upper 10% of the phenotypic distribution for a trait with h^2 = 0.1, 71 progeny must be tested to be 80% certain of selecting one or more progeny from the upper 10% of the genotypic distribution, while 498 progeny must be tested to be 80% certain of selecting one or more progeny from the upper 1% of the genotypic distribution (Table 1). The former requires selecting 7 of 71 progeny, whereas the latter requires selecting 50 of 498 progeny (Table 1). Chances are that none of the geno-

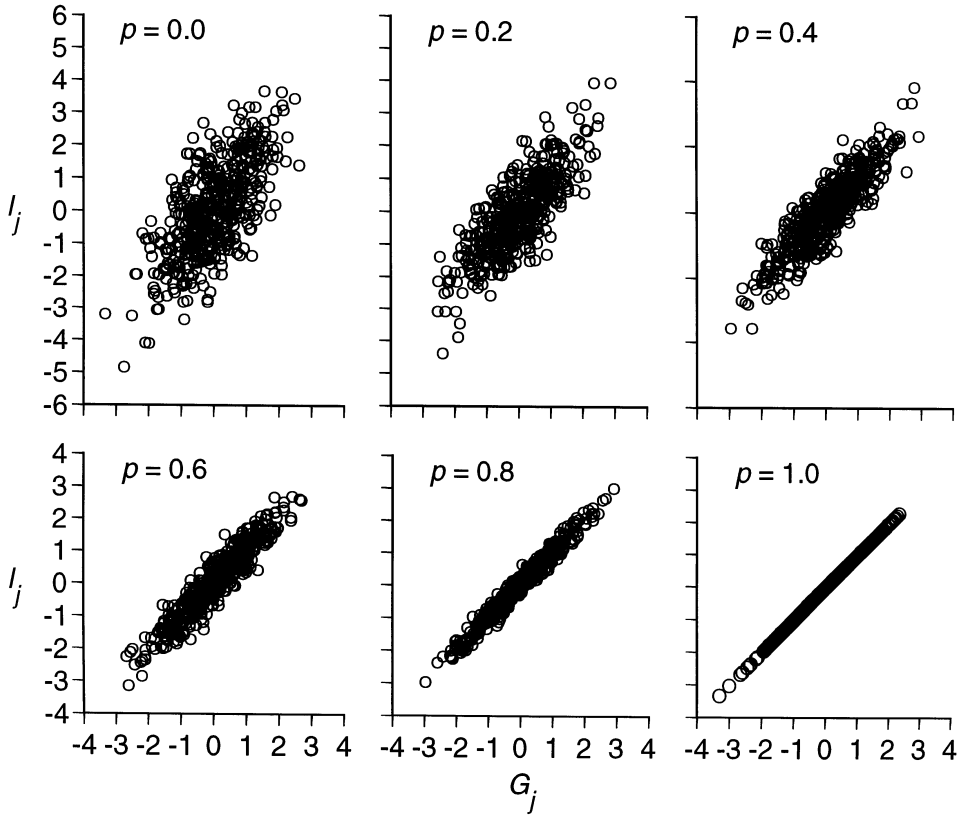

Fig. 6. Genotypic (G_j) and MAS index (I_j) effects for samples of 500 progeny simulated for a trait with $h^2 = 0.5$ for phenotypic selection ($p = 0$) and marker-assisted selection ($p = 0.2$ to 1.0 by 0.2). The estimated heritabilities ($\hat{h}^2$) for the samples simulated using $p = 0.0$, 0.2, 0.4, 0.6, 0.8, and 1.0 samples were 0.475, 0.668, 0.798, 0.859, 0.974, and 1.0, respectively.

types in the first sample (7 of 71) would have genotypic means $2.326\sigma_G$ greater than μ ($Pr_{MAS} \cdot 7 = 0.032 \cdot 7 = 0.224$), but one or more ($0.224 \cdot 7 = 1.568$) should have genotypic means greater means $1.282\sigma_G$ greater than μ. The second sample of selected progeny (50 of 498) should have 11.2 ($0.224 \cdot 50$) progeny with genotypic means $1.282\sigma_G$ greater than μ and 1.6 ($0.032 \cdot 50$) progeny with genotypic means $2.326\sigma_G$ greater than μ; thus, the genotypes of 49 of 50 of the selected progeny (98%) would be inferior when the goal is to select a line from the upper 1% of the genotypic distribution.

DISCUSSION

Lande and Thompson (1990) used $E = R_{MAS}/R_{PS}$ to predict the efficiency of MAS for individual or family selection, where R_{MAS} is the gain from marker-assisted selection and R_{PS} is the gain from phenotypic selection. The efficiency measure described in this paper (E_c) differs in a few ways from E. E predicts efficiency for very large samples, whereas E_c predicts efficiency for finite samples. More specifically, E_c predicts how MAS affects n (the minimum number of progeny a breeder needs to test to be assured of selecting one or more superior genotypes). E_c is affected by i, g, and the i by g interaction (Fig. 3), whereas E is not (Lande and Thompson, 1990) (E is only affected by selection intensity when different selection intensities are used for phenotypic and marker-assisted selection).

The underlying differences between E_c and E are important because E_c is a more optimistic predictor of efficiency than E for some i, g, p, and h^2. Lande and Thompson (1990) reported efficiencies between 1.0 and 3.2 for h^2 from 0.1 to 1.0 and p from 0.0 to 1.0 for individual selection (or family mean selection with large families). E_c ranged from 1.0 to 16.7 for h^2 from 0.1 to 1.0 and p from 0.0 to 1.0 (Fig. 6) and was greater than E for most i, g, p, and h^2. E predicts that MAS is not cost effective when the cost of MAS is one to three times more than the cost of phenotypic selection, whereas E_c predicts that MAS is not cost effective when the cost of MAS is one to 17 times more than the cost of phenotypic selection for h^2 from 0.1 to 1.0.

The efficiencies predicted by E_c or any other theoretical estimate of efficiency overestimate the true efficiency gained by MAS when the estimated marker effects ($\hat{p}$) and true QTL effects (p) are not perfectly correlated. This is the reality in practice and has many causes (Beavis et al., 1991, 1994; Beavis and Smith, 1996; Beavis, 1994, 1997; Bulmer, 1971; Churchill and Doerge, 1994; Davarsi and Soller, 1994, 1995; Davarsi et al., 1993; Doerge et al., 1994; Gimelfarb and Lande, 1994a, 1995; Jansen, 1993; Jansen and Stam, 1994; Knapp et al., 1993; Knapp, 1994b; Stuber and Sisco, 1991; Stuber, 1994, 1995; Weller, 1993; Xu and Atcheley, 1995; Xu, 1996; Visscher et al., 1996; Zeng, 1994; Zhang and Smith, 1992, 1993). First, significant marker effects could be false positives. Putting selection pressure on markers segregating independent of QTL (false positives) increases genetic drift and erodes genetic variance (Bulmer, 1971; Gimelfarb and Lande, 1995; Zhang and Smith, 1992, 1993). Second, non-significant marker ef-

Table 1. The number of progeny a breeder needs to test, n_{PS} for phenotypic selection ($p = 0$) and n_{MAS} for marker-assisted selection ($p > 0$), to be $100(1 - c)$% certain of selecting one or more progeny with genotypic values greater than g' among progeny with phenotypic or index values greater than x' or i' for a normally distributed quantitative trait. n_{PS} and n_{MAS} were calculated using factorial combinations of genotypic selection thresholds (g), phenotypic or index selection thresholds (x or i), heritabilities (h^2 or h_I^2), and assurance probabilities ($1 - c$), where $g = (g' - \mu)/\sigma_G$, $x = (x' - \mu)/\sigma_P$, $i = (i - \mu)/\sigma_I$, μ is the population mean, σ_G^2 is the genotypic variance, σ_P^2 is the phenotypic variance, and σ_I^2 is the index variance.

			$1 - c$				
x or i	g	h^2 or h_I^2	0.80	0.90	0.95	0.99	0.999
1.282	1.282	0.1	71	102	132	203	305
		0.2	54	78	101	155	233
		0.3	45	64	83	127	192
		0.4	38	55	71	109	163
		0.5	33	47	62	95	142
		0.6	29	42	54	84	125
		0.7	26	37	48	74	111
		0.8	23	33	43	65	98
		0.9	20	28	37	57	86
		1.0	15	22	28	44	66
	1.645	0.1	125	179	233	359	538
		0.2	92	132	172	264	396
		0.3	74	106	138	212	319
		0.4	62	89	116	179	267
		0.5	54	77	100	154	231
		0.6	47	67	88	135	202
		0.7	42	60	78	119	179
		0.8	37	53	69	106	159
		0.9	33	48	62	95	143
		1.0	31	45	58	90	135
	2.326	0.1	498	712	927	1425	2137
		0.2	349	499	650	998	1498
		0.3	276	395	513	789	1184
		0.4	232	331	431	663	994
		0.5	202	290	377	579	869
		0.6	183	261	340	522	784
		0.7	170	243	316	485	728
		0.8	162	232	302	465	697
		0.9	160	229	298	458	688
		1.0	161	229	299	459	688
1.645	1.645	0.1	214	307	399	614	920
		0.2	150	215	279	429	644
		0.3	117	167	217	334	501
		0.4	95	136	177	273	409
		0.5	80	114	149	228	343
		0.6	68	98	127	195	293
		0.7	59	84	109	168	252
		0.8	50	72	94	144	217
		0.9	43	61	80	122	184
		1.0	31	45	58	90	135
	2.326	0.1	803	1149	1495	2298	3448
		0.2	522	746	971	1492	2239
		0.3	389	556	723	1112	1668
		0.4	309	442	576	885	1328
		0.5	257	367	478	735	1102
		0.6	219	314	408	628	942
		0.7	192	275	358	550	825
		0.8	173	247	322	495	742
		0.9	162	232	301	463	695
		1.0	160	229	297	457	686
2.326	2.326	0.1	2683	3839	4994	7677	11 515
		0.2	1530	2189	2848	4378	6567
		0.3	1038	1485	1931	2969	4454
		0.4	763	1092	1420	2183	3275
		0.5	588	841	1093	1682	2522
		0.6	466	666	867	1332	1998
		0.7	375	536	697	1072	1608
		0.8	303	433	563	866	1299
		0.9	241	345	449	689	1034
		1.0	160	229	298	458	687

Table 2. The number of progeny a breeder needs to test, n_{PS} for phenotypic selection ($p = 0$) and n_{MAS} for marker-assisted selection ($p > 0$), to be 99% certain ($1 - c = 0.99$) of selecting one or more progeny with genotypic values greater than g' among progeny with index values greater than i for a normally distributed quantitative trait. n_{PS} and n_{MAS} were calculated using factorial combinations of genotypic selection thresholds (g), index selection thresholds (i), heritabilities (h^2), where $g = (g' - \mu)/\sigma_G$, $i = (i - \mu)/\sigma_I$, and μ is the population mean, σ_G^2 is the genotypic variance, and σ_I^2 is the index variance.

			p					
i	g	h^2	0.0	0.2	0.4	0.6	0.8	1.0
1.282	1.282	0.1	203	68	55	50	46	44
		0.2	155	77	61	53	47	44
		0.5	95	79	67	58	51	44
		0.9	57	57	56	55	52	44
	1.645	0.1	359	109	93	90	90	90
		0.2	264	124	100	91	90	90
		0.5	154	128	109	97	90	90
		0.9	95	95	94	93	91	90
	2.326	0.1	1424	468	458	458	458	458
		0.2	998	495	459	458	458	458
		0.5	579	503	468	459	458	458
		0.9	458	458	458	458	458	458
1.645	1.645	0.1	614	150	118	104	96	90
		0.2	429	177	132	111	99	90
		0.5	229	183	150	126	106	90
		0.9	122	121	119	116	110	90
	2.236	0.1	463	462	461	460	458	458
		0.2	463	462	461	460	458	458
		0.5	463	462	461	460	458	458
		0.9	463	462	461	460	458	458
2.326	2.326	0.1	7677	916	655	554	497	458
		0.2	4378	1151	762	605	517	458
		0.5	1682	1208	914	714	569	458
		0.9	689	681	668	645	600	458

fects could be false negatives missed because of low statistical power. Power can be increased by increasing n and other experimental variables and is affected by the estimation or model fitting procedure (Jansen and Stam, 1994; Knapp and Bridges, 1990; Knapp et al., 1993; Soller and Beckmann, 1990; Van Oijen, 1992; Zeng, 1994). Third, when individual markers are selected (as opposed to flanking markers), crossovers can arise between the selected markers and the underlying QTL, thereby reducing some of the predicted gain (Gimelfarb and Lande, 1995). Fourth, when markers flanking a "QTL LOD peak" are selected, the QTL might not reside between the selected markers. If so, crossovers can arise between the selected markers and the underlying QTL and reduce some of the predicted gain. Fifth, a QTL could be a block of linked genes dispersed between and around one or more selected markers. These linkage blocks can undergo recombination, thereby producing genotypes fixed for favorable and unfavorable alleles and reducing the predicted gain. Sixth, sampling biases for unmapped QTL can produce misleading estimates of QTL effects, some of which lead to false positives or false negatives (Knapp et al., 1993). Finally, because markers linked to QTL can be rapidly fixed, disequilibria are often produced between selected QTL and unselected QTL underlying the residual polygenic variance (Bulmer, 1971).

Despite these pitfalls, many of the QTL reported in

empirical studies, particularly among progeny from crosses between inbred lines, are undoubtedly bona fide (Beavis et al., 1991, 1994; Beavis and Smith, 1996; Dudley, 1993; Hayes et al., 1996; McCouch and Doerge, 1995; Oziel et al., 1996; Stuber, 1992, 1994, 1995; Stuber and Sisco, 1991; Stuber et al., 1992; Tanskley, 1993; Tanksley et al., 1989). The most efficient strategy should be to use very stringent significance thresholds (Doerge et al., 1993), in addition to rigorously estimating the parameters using multilocus methods for linked and unlinked QTL (Jansen and Stam, 1994; Knapp et al., 1993; Knapp, 1994b; Martinez and Curnow, 1992; Visscher et al., 1996; Zeng, 1994). The goal is to produce the most accurate estimate of p for a particular data set, even if p falls well short of 1.0, so that selection pressure is only put on bona fide QTL. Knapp and Bridges (1990) showed that $\hat{p} = 1$ can be achieved with any balanced data set and many unbalanced data sets by randomly selecting markers until the between progeny degrees of freedom are used up. This can only be done when there are more marker genotypes than progeny, which is the predominant situation in practice. The selected markers need not even be linked to QTL to get $\hat{p} = 1$. Therein lies the problem: misleading QTL and σ^2_M estimates can be produced, even when the analysis has been rigorous and the statistics show that the experimenter has accounted for most or all of the genetic variance in the test sample.

Putting selection pressure only on those markers with highly significant additive effects has merit because even modest increases in p greatly increase h^2_I, Pr_{MAS}, and E_c and decrease n_{MAS} (Tables 1 and 2 and Fig. 1–3). Additionally, diminishing returns are produced by adding more and more significant markers to the index (increasing $\hat{p}$) when h^2 is low. Using very high significance thresholds to select markers for the index should increase the accuracy of the index and still produce significant gains over phenotypic selection alone.

Using an index based on only the most important QTL may seem illogical because there is a tendency to assume that QTL with large effects are bound to be fixed anyway, with or without MAS; however, the most outstanding genotypes for low heritability traits are often not present in the selected fraction because their phenotypic scores are below the selection cutoff (Fig. 4–6), even though many are homozygous or heterozygous for favorable alleles for large effect QTL. This problem is exacerbated with small samples (Tables 2 and 3; Beavis, 1994, 1997; Gimelfarb and Lande, 1994a, 1994b, 1995; Zhang and Smith, 1992, 1993).

Pr_{MAS}, E_n, and n_{MAS} predict that exceptional gains can be produced by MAS under certain circumstances (Fig. 1–3 and Tables 1 and 2). First, MAS is most cost effective when the goal is to select a genotype from the upper 1 to 2% of the genotypic distribution for most quantitative traits. Second, once the selection goal is set, the selection intensity must be carefully chosen to gain the most efficiency. Setting $i = g$ produces the greatest efficiency (Fig. 3). Setting i less than g is ill advised for some h^2 and p, but not others (Fig. 2 and 3). Setting i greater than g produces the fastest increases in Pr_{MAS} up to the maximum Pr_{MAS} for a specific h^2 and g (Fig. 2), but is ill advised from an efficiency standpoint (Fig. 3). Third, resources should be concentrated on the most promising crosses if there is a sound basis for choosing crosses. There is more merit to testing a large number of progeny per cross from a small number of crosses as opposed to a small number of progeny per cross from a large number crosses, primarily because the accuracy of QTL and MAS index parameter estimates increases and the probability of selecting an outstanding genotype increases as n increases (n decreases as the number of crosses increases for a fixed number of experimental units). Choosing crosses, however, is not a trivial problem. Methods have been developed for choosing parents for crosses for enhancing the parents of single-cross hybrids (Dudley, 1984, 1987; Gerloff and Smith, 1988a,b); however, much less work has been done on this problem in self-pollinated crops (Panter and Allen, 1995a,b; St. Martin et al., 1996).

REFERENCES

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250–266. *In* Proc. 49th Annu. Corn Sorghum Indus. Res. Conf. Am. Seed Trade Assoc., Washington, DC.

Beavis, W.D. 1997. QTL analyses: power, precision, and accuracy. p. 145–162. *In* A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Beavis, W.D., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with quantitative genetic loci. Theor. Appl. Genet. 83:141–145.

Beavis, W.D., O.S. Smith, D. Grant, and R. Fincher. 1994. Identification of quantitative trait loci using a small sample of topcrossed and F_4 progeny from maize. Crop Sci. 34:882–896.

Beavis, W.D., and S. Smith. 1996. Molecular marker assisted breeding in a company environment. p. 259–272. *In* B.W.S. Sobral (ed.) The impact of plant molecular genetics. Birkhaeuser, Boston.

Bulmer, M.G. 1971. The effect of selection on genetic variability. Am. Nat. 108:45–58.

Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genet. 138:963–971.

Davarsi, A., and M. Soller. 1994. Optimum spacing of genetic markers for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet. 89:351–357.

Davarsi, A., and M. Soller. 1995. Advanced intercross lines: An experimental population for fine genetic mapping. Genet. 141:1199–1207.

Davarsi, A., A. Weinreb, V. Minke, J.I. Weller, and M. Soller. 1993. Detecting marker-QTL linkage and estimating QTL gene effects and map location using a saturated genetic map. Genet. 134: 943–951.

Doerge, R.W., Z.-B. Zeng, and B.S. Weir. 1994. Statistical issues in the search for genes affecting quantitative traits in populations. p. 15–26. *In* Proc. Second Plant Breed. Symp., CSSA and ASHS, Corvallis, OR.

Dudley, J.W. 1984. A method of identifying lines for use in improving parents of a single-cross. Crop Sci. 24:355–357.

Dudley, J.W. 1987. Modification of methods for identifying inbred lines useful for improving parents of elite single crosses. Crop Sci. 27:944–947.

Dudley, J.W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33:660–668.

Falconer, D.S. 1981. Introduction to quantitative genetics. Longman Group Ltd., Essex.

Geiger, H.E. 1984. Selection in early generations. p. 65–71. *In* W. Lange et al. (ed.) Efficiency in plant breeding. Proc. 10th EUCARPIA Congress, Wageningen, the Netherlands.

Gerloff, J., and O.S. Smith. 1988a. Choice of method for identifying germplasm with superior alleles. I. Theoretical results. Theor. Appl. Genet. 76:209–216.

Gerloff, J., and O.S. Smith. 1988b. Choice of method for identifying germplasm with superior alleles. I. Computer simulation results. Theor. Appl. Genet. 76:217–227.
Gimelfarb, A., and R. Lande. 1994a. Simulation of marker assisted selection for non-additive traits. Genet. Res. Camb. 64:127–136.
Gimelfarb, A., and R. Lande. 1994b. Simulation of marker assisted selection in hybrid populations. Genet. Res. Camb. 63:39–47.
Gimelfarb, A., and R. Lande. 1995. Marker-assisted selection and marker-QTL associations in hybrid populations. Theor. Appl. Genet. 91:522–528.
Hallauer, A.R., and F. Miranda. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames.
Hayes, P.M., F.Q. Chen, A. Kleinhofs, A. Kilian, and D.E. Mather. 1996. Barley genome mapping and its applications. p. 229–249. *In* P.P. Jauhar (ed.) Methods of genome analysis in plants. CRC Press, Boca Raton, FL.
Hazel, L.N. 1943. The genetic basis for constructing selection indices. Genet. 28:476–490.
Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci. Genet. 135:205–211.
Jansen, R.C., and P. Stam. 1994. High resolution of quantitative traits into multiple loci via interval mapping. Genet. 136:1447–1455.
Johnson, B. 1989. The probability of selecting genetically superior S_2 lines from a maize population. Maydica 34:5–14.
Knapp, S.J. 1994a. Selection using molecular marker indexes. p. 1–8. *In* Proc. Second Plant Breed. Symp., CSSA and ASHS, Corvallis, OR.
Knapp, S.J. 1994b. Mapping quantitative trait loci. p. 58–96. *In* R.L. Phillips and I.K. Vasil (ed.) DNA markers in plants. Kluwer, Dordrecht, the Netherlands.
Knapp, S.J., and W.C. Bridges. 1990. Using molecular markers to estimate quantitative trait locus parameters: Power and genetic variances for unreplicated and replicated progeny. Genet. 126: 769–777.
Knapp, S.J., W.C. Bridges, and B.-H. Liu. 1993. Mapping quantitative trait loci using nonsimultaneous and simultaneous estimators and hypothesis tests. p. 209–235. *In* J.S. Beckmann and T.S. Osborn (ed.) Plant genomes: Methods for genetic and physical mapping. Kluwer, Dordrecht, the Netherlands.
Lande, R. 1992. Marker-assisted selection in relation to traditional methods of plant breeding. p. 437–451. *In* H.T. Stalker and J.P. Murphy (ed.) Plant breeding in the 1990s. Proc. Symp. Plant Breeding in the 1990s, C.A.B. International, Wallingford, UK.
Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genet. 124: 743–756.
Lin, C.Y., and F.R. Allaire. 1977. Heritability of a linear combination of traits. Theor. Appl. Genet. 51:1–3.
Martinez, O., and R.N. Curnow. 1992. Estimation of the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. 85:480–488.
McCouch, S.R., and R.W. Doerge. 1995. QTL mapping in rice. Trends Genet. 11:482–487.
Oziel, A., P.M. Hayes, F.Q. Chen, and B. Jones. 1996. Application of quantitative trait locus mapping to the development of winter-habit malting barley. Plant Breed. 115:43–51.
Panter, D.M., and F.L. Allen. 1995. Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. Crop Sci. 35:405–410.
Panter, D.M., and F.L. Allen. 1995. Using best linear unbiased predictions to enhance breeding for yield in soybean: II. Selection of superior crosses from a limited number of yield trials. Crop Sci. 35:410–414.
Perlin, M.W., G. Lancia, N. See-Kiong. 1995. Toward fully automated genotyping: Genotyping microsattelite markers by deconvolution. Am. J. Hum. Genet. 57:1199–1210.
Robson, D.S., L.Powers, and N.S. Urquhart. 1967. The proportion of genetic deviates in the tails of a normal distribution. Der Zuechter 37:205–216.
SAS. 1985. SAS/IML user's guide. SAS Inst., Cary, NC.
Schwengel, D.A. A.E. Jedlicka, E.J. Nanthakumar, J.L. Weber, and R.C. Levitt. 1994. Comparison of fluorescence-based semi-automated genotyping of multiple microsattelite loci with autoradiographic techniques. Genomics 22:46–54.
Soller, M., and J.S. Beckmann. 1990. Marker-based mapping of quantitative trait loci using replicated progenies. Theor. Appl. Genet. 80:205–208.
Smith, H.F. 1936. A discriminant function for plant selection. Ann. Eugenics 7:240–250.
Snape, J.W., and E. Simpson. 1984. Early generation selection and rapid generation advancement methods in autogamous crops. *In* W. Lange et al. (ed.) Efficiency in plant breeding. Proc. 10Th EUCARPIA Congress, Wageningen, the Netherlands.
Sneep, J. 1977. Selection for yield in early generations of self-fertilizing crops. Euphytica 26:27–30.
Sneep, J. 1984. Is selection on quantitative characters between F_3 lines in small grains feasible? *In* W. Lange et al. (ed.) Efficiency in plant breeding. Proc. 10th EUCARPIA Congress, Wageningen, the Netherlands.
St. Martin, S.K., K.S. Lewers, R.G. Palmer, and B.R. Hedges. 1996. A testcross procedure for selecting exotic strains to improve pure-line cultivars in predominantly self-fertilized crops. Theor. Appl. Genet. 92:78–82.
Stromberg, L.D., J.W. Dudley, and G.K. Rufener. 1994. Comparing conventional early generation selection with molecular marker assisted selection in maize. Crop Sci. 34:1221–1225.
Stuber, C.W. 1992. Biochemical and molecular markers in plant breeding. Plant Breed. Res. 9:37–61.
Stuber, C.W.. 1994. Successes in the use of molecular markers for yield enhancement in corn. p. 232–238. *In* Proc. 49th Annu. Corn Sorghum Indus. Res. Conf. Am. Seed Trade Assoc., Washington, DC.
Stuber, C.W. 1995. Mapping and manipulating quantitative traits in maize. Trends Genet. 11:477–481.
Stuber, C.W., and P.H. Sisco. 1991. Marker facilitated transfer of QTL alleles between elite inbred lines and responses in hybrids. p. 104–113. *In* Proc. 46th Annu. Corn Sorghum Indus. Res. Conf. Am. Seed Trade Assoc., Washington, DC.
Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genet. 132:823–839.
Tanksley, S.D. 1993. Mapping polygenes. Annu. Rev. Genet. 27: 205–234.
Tanksley, S.D., N.D. Young, A.H. Paterson, M.W. Bonierbale. 1989. RFLP mapping in plant breeding: New tools for an old science. Biotechnol. 7:257–264.
Van Ooijen, J.W. 1992. Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. 84:803–811.
Visscher, P.M., C.S. Haley, and R. Thompson. 1996. Marker-assisted introgression in backcross breeding programs. Genet. 144:1923–1932.
Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1996. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407–4414.
Weber, W.E. 1984. Selection in early generations. *In* W. Lange et al. (ed.) Efficiency in plant breeding. Proc. 10th EUCARPIA Congress, Wageningen, the Netherlands.
Weller, J.I. 1993. Statistical methodologies for mapping and analysis of quantitative trait loci. p. 181–207. *In* J.S. Beckmann and T.S. Osborn (ed.) Plant genomes: Methods for genetic and physical mapping. Kluwer, Dordrecht, the Netherlands.
Wolfram, S. 1989. Mathematica: A system for doing mathematics by computer. Addison-Wesley, Redwood City, CA.
Wricke, G., and W.E. Weber. 1986. Quantitative genetics and selection in plant breeding. Walter de Gruyter, New York.
Xu, S. 1996. Computation of the full likelihood function for estimating variance at a quantitative trait locus. Genet. 144:1951–1960.
Xu, S., and W.R. Atcheley. 1995. A random model approach to interval mapping of quantitative trait loci. Genet. 141:1189–1197.
Zeng, Z.-B. 1994. Precision mapping of quantitative trait loci. Genet. 136:1457–1468.
Zhang, W., and C. Smith. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. Theor. Appl. Genet. 83:813–820.
Zhang, W., and C. Smith. 1993. Simulation of marker-assisted selection utilizing linkage disequilibrium: The effects of several additional factors. Theor. Appl. Genet. 86:492–496.