14

Long-term Selection with Known Quantitative Trait Loci

*Jack C. M. Dekkers and Petek Settar**
Department of Animal Science
Iowa State University
Ames, Iowa 50011-3150

I. INTRODUCTION

Most traits of interest in agriculturally important species are quantitative and affected by a potentially large numbers of genes. Until recently, most artificial selection for quantitative traits, both for agronomic and experimental purposes, has been based on observed phenotype or on

*The authors gratefully acknowledge contributions to the work presented in this manuscript by Reena Chakraborty, Rohan Fernando, and Laurence Moreau. Discussions with Bill Muir are acknowledged also. Aspects of the work presented herein were financially supported by SYGEN/PIC, the NRI Competitive Grants Program/USDA (award no. 98-35205-6736) and supported by Hatch Act and State of Iowa Funds through the Iowa Agriculture and Home Economics Experiment Station, Ames (Project No. 3456).

Plant Breeding Reviews: Part 1: Long-term Selection: Maize, Volume 24 Edited by Jules Janick
 ISBN: 978-0-471-35316-4

estimates of breeding values derived from phenotype. In this quantitative genetic approach (Falconer and Mackay 1996), efficient selection strategies are designed based on population parameters for the collective effect of all genes that affect the trait (e.g., heritability) without knowledge of the number of genes that affect the trait, nor of their individual effects and frequencies. Implications for long-term response to selection on phenotype and for the design of selection strategies to maximize long-term response were reviewed by others in this symposium.

Advances in molecular genetics have enabled detection of at least some of the genes or genomic regions that affect quantitative traits, so-called quantitative trait loci (QTL). Strategies and advances for the detection of QTL were reviewed by Mauricio (2001) for plants and by Anderson (2001) for livestock. Dekkers and Hospital (2002) reviewed strategies and challenges for the use of molecular information to enhance genetic improvement through marker-assisted selection (MAS).

Knowledge of the genotype of individuals for QTL, either directly or through linked genetic markers, can enhance genetic improvement of quantitative traits because of increased selection accuracy and/or reduced generation intervals, along with potentially reduced requirements for phenotypic information. The main focus of studies on MAS, which primarily have been theoretical or based on computer simulations, has been on enhancing short- to medium-term responses to selection. The objective of this paper is to review and investigate the use of QTL information to enhance long-term response to selection for a quantitative trait. The main focus will be on traits for which phenotype is observed prior to selection and on cases where the QTL can be genotyped directly, rather than indirectly through linked genetic markers. For the most part, however, the concepts developed and discussed herein also apply to selection on linked markers. For illustrative purposes, a completely additive genetic model is assumed throughout.

II. STANDARD STRATEGIES FOR MARKER-ASSISTED SELECTION

A. Combining Phenotype and QTL Information

When distinguishing QTL that have been mapped from other background genes that affect the trait, which will be referred to as polygenes, the genetic value g_i of an individual i, can be partitioned into the sum of genetic values at the QTL, g_{Qi}, and the sum of genetic values at polygenes, g_{pi}: $g_i = g_{Qi} + g_{pi}$. Molecular genetic information provides

information that can be used to estimate g_{Qi}, whereas an individual's phenotype provides information on the collective effect of all genes. Unless all QTL that affect the trait have been identified, selection on QTL must be combined with selection on phenotypic information, to ensure simultaneous improvement of both g_{Qi} and g_{pi}. Lande and Thompson (1990) suggested that QTL and phenotypic information should be combined in an index of the following form:

$$I_i = b_Q \hat{g}_{Qi} + b_P P_i$$

where $\hat{g}_{Qi}$ is the molecular score for individual i, that is, the individual's estimated breeding value for the QTL, P_i is the individual's phenotype, and b_Q and b_P are index weights. The molecular score, $\hat{g}_{Qi}$, can be computed as the sum over QTL or markers of estimates of effects on phenotype based on the individual's QTL or marker genotypes. Lande and Thompson (1990) showed that index weights could be derived by standard selection index theory, given the proportion of genetic variance explained by the QTL or markers (r), and the heritability of the trait (h^2), resulting in the following index weights:

$$b_Q = \frac{1 - h^2}{1 - rh^2} \quad \text{and} \quad b_p = h^2 \frac{1 - r}{1 - rh^2}$$

It is useful to note that this index can be reparameterized into an equivalent index of molecular score and phenotype adjusted for the molecular score as follows:

$$I_i' = b_Q' \hat{g}_{Q\,i} + b_P' P_i'$$

Where $P_i' = P_i - \hat{g}_{Qi}$. Using selection index theory and defining polygenic heritability as the heritability of phenotype adjusted for molecular score:

$$h_p^2 = \frac{h^2(1 - r)}{1 - rh^2}$$

weights for this index can then be derived to be independent of r and equal to: $b_Q' = 1$ and $b_P' = h_p^2$. Thus, the resulting index is: $I_i' = g_{Qi} + h_p^2 P_i'$. Note that the second term in this index, $h_p^2 P_i'$, represents the individual's estimated breeding value for polygenes, $\hat{g}_{pi}$. Thus, this index can be represented in a more general form as:

$$I_i' = \hat{g}_{Qi} + \hat{g}_{pi}$$

This formulation of the index is easily extended to situations in which indexes of phenotypes of relatives are used to estimate the breeding value for polygenic effects. These extensions include the marker-assisted best linear unbiased prediction models to derive estimated breeding values (EBV) with marker information that were developed by Fernando and Grossman (1989). Another advantage of index I' over index I is that its index weights remain constant over generations, whereas weights for index I must be updated each generation as the frequency of the QTL, and therefore r, changes.

It is useful to note that phenotypic selection can also be written as selection on an index of breeding values for the QTL and polygenes by noting that selection on P_i is equivalent to selection on $h_p^2 P_i$, which can be written as

$$h_p^2 P_i = h_p^2 g_{Q_i} + h_p^2 P_i' = h_p^2 g_{Q_i} + \hat{g}_{pi}$$

Thus, with phenotypic selection, the emphasis on the QTL relative to the EBV for polygenes is equal to the polygenic heritability, h_p^2, instead of 1 as in MAS.

B. Short-term Response to Marker-assisted Selection

In theory, for an additive QTL and if parameters are known without error, selection on the Lande and Thompson (1990) index for MAS, which will be referred to as standard MAS, maximizes the average response to selection over a single generation. Using selection index theory, Lande and Thompson (1990) showed that response from MAS relative to selection on phenotype alone is expected to equal $\sqrt{\frac{r}{h^2} + \frac{(1-r)^2}{1-rh^2}}$. This equation shows that MAS will be most beneficial for traits with low heritability and when the molecular score explains a large proportion of the genetic variance. The ability to increase response to selection in the short- and medium-term using selection criteria similar to those presented above has been demonstrated by computer simulation in several studies (Zhang and Smith 1992; Gimelfarb and Lande 1994; Meuwissen and Goddard 1996). Sustained breeding programs must, however, also consider the longer-term consequences of selection, which will be discussed in the following.

C. Long-term Response to Marker-assisted Selection

Gibson (1994) was the first to show that, although standard MAS can increase response to selection in the short-term, it can result in lower cumulative responses in the longer term than selection on phenotype alone. Gibson (1994) used a deterministic simulation model for selection on a biallelic QTL (alleles b and B) of known effect (genetic values of $-a$, 0, and $+a$ for bb, Bb, and BB genotypes), by modeling selection across three normal distributions (Fig. 14.1). For standard MAS, each distribution represents the distribution of index I' for a given genotype, with mean g_Q (= $-a$, 0, and $+a$) and standard deviation $h_p^2\sigma_p'$, where σ_p' is the

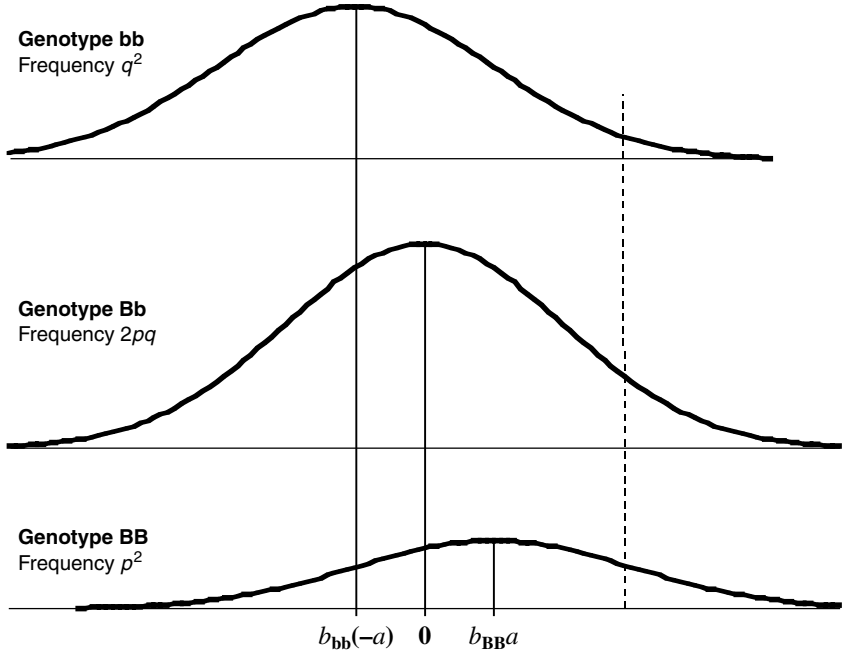

Fig. 14.1. Deterministic model for selection in a population of infinite size on an index of the genetic value of a QTL (for genotypes BB, Bb, and bb) and an estimated breeding value of phenotype adjusted for g_Q, $h_p^2P_i'$, where P_i' is the adjusted phenotype, with heritability h_p^2: $I_i' = b_Q g_{Qi} + h_p^2 P_i'$, where b_Q is the index weight, which is equal to 1 for standard marker-assisted selection, equal to h_p^2 for phenotypic selection, and optimized for optimal marker-assisted selection. The three distributions represent the distribution of index I_i' for each genotype, with mean $b_Q g_Q$ and standard deviation equal to = $\sqrt{Var(h_p^2 P_i')} = h_p^2 \times$ pheno-typic standard deviation. The broken line represents the truncation point for selection on the index.

standard deviation of P'. Note that the selection model depicted in Fig. 14.1 also holds for phenotypic selection, except that means of the distributions are $-h_p^2 a$, 0, and $+h_p^2 a$, according to the implicit index for phenotypic selection: $h_p^2 P_i = h_p^2 g_{Qi} + h_p^2 P_i'$.

Fig. 14.2 illustrates responses to selection on phenotype and to standard MAS based on the deterministic model depicted in Fig. 14.1 for an example situation. For illustrative purposes, the example reflects a QTL of very large effect (the difference between homozygotes is $2a = 1.5\sigma_p'$). Similar trends are observed for QTL of smaller effect, although the differences between phenotypic selection and MAS are smaller. Fig. 14.2 clearly shows the extra response from MAS during early generations. By generation 5, however, cumulative response from phenotypic selection exceeds that from MAS. As expected, MAS fixes the QTL at a faster rate than phenotypic selection (Fig. 14.2b). The increased selection emphasis on the QTL, however, results in lower response in polygenes (Fig. 14.2c). Although polygenic response per generation returns to maxi-

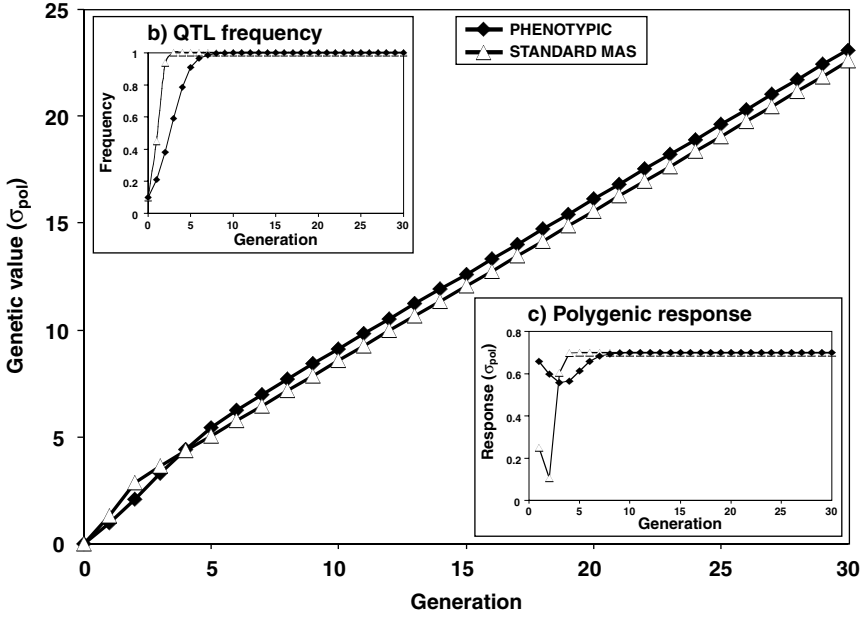

Fig. 14.2. Responses to standard MAS and phenotypic selection based on the deterministic model in a population of infinite size. Selection is of the top 20% of males and females for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. The main graph shows cumulative total response to selection, expressed in polygenic standard deviations (σ_{pol}); b) frequency of the favorable QTL allele; and c) polygenic response per generation.

mum as soon as the QTL is fixed, that is, sooner for MAS than for phenotypic selection, the extra polygenic response that is lost in early generations with MAS is never regained in later generations, which is the reason for the lower cumulative response for MAS in the longer term.

Results illustrated in Fig. 14.2 are based on several simplifying assumptions for the polygenic component of the genetic model: (a) the infinitesimal model for polygenes, that is, an infinite number of polygenes of small effect; (b) large population size, that is, no inbreeding or drift; and (c) genetic variance contributed by polygenes remains constant over generations, that is, no gametic phase disequilibrium among polygenes (Bulmer 1980). The deterministic model does account for the gametic phase disequilibrium between the QTL and polygenes that is induced by simultaneous selection on the QTL and polygenes (Dekkers and van Arendonk 1998). This is reflected in a negative association between the QTL and polygenes, such that individuals with a (un)favorable QTL genotype tend to have poorer (better) polygenic breeding values. The creation of this negative association by selection is illustrated in Fig. 14.1 by noting that individuals with a BB genotype are less intensely selected for polygenes than individuals with a bb genotype for the QTL. A negative association is created by both phenotypic selection and MAS but is larger for MAS because of the greater emphasis on the QTL (Dekkers and van Arendonk 1998).

Despite the simplifying assumptions of the deterministic model, the results illustrated in Fig. 14.2 have been repeated in several studies by stochastic simulation (e.g., Larzul et al. 1997; Pong-Wong and Woolliams 1998). A stochastic model simulates individuals in the population under selection, rather than population distributions, and does not require many of the assumptions that are inherent to the deterministic model depicted in Fig. 14.1. Typical results from such stochastic simulations are demonstrated in Fig. 14.3, which represents the results of simulating selection in a population of 250 males and 250 females, with 20% selected for each sex. Three different genetic models were used for polygenes: the infinitesimal genetic model and models in which the polygenic component is simulated by 50 or 10 individual loci. Results for the stochastic model were averaged over 500 replicate simulations.

Fig. 14.3 focuses on the difference in cumulative responses between MAS and phenotypic selection over generations, rather than the absolute responses illustrated in Fig. 14.2. For a given method of selection (MAS or phenotypic selection), absolute cumulative responses to selection (not shown) differed between genetic models; responses were greatest for the deterministic model, followed by the infinitesimal model, and the finite locus models with 50 and 10 polygenes. Average rates of change

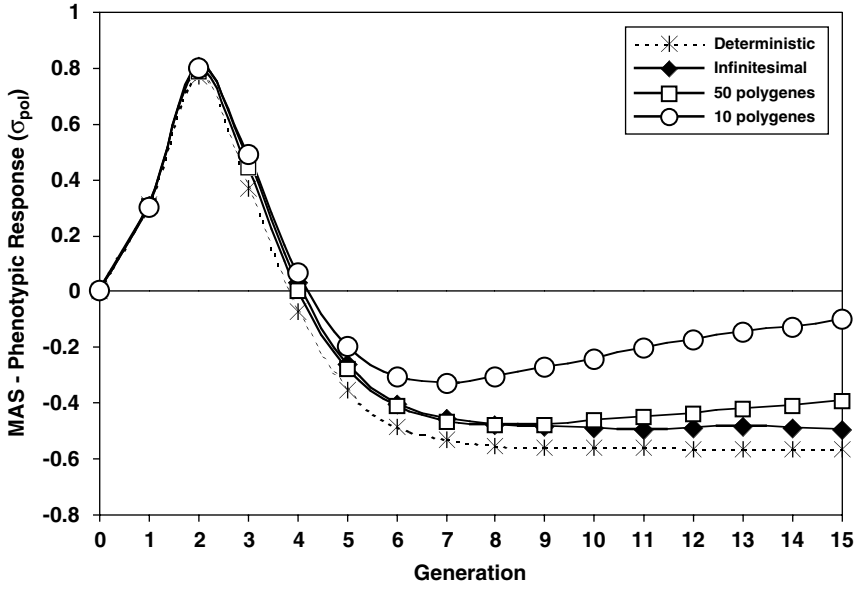

Fig. 14.3. Cumulative responses to standard MAS as a deviation from cumulative response for phenotypic selection (MAS response – phenotypic response, expressed in polygenic standard deviations, σ_{pol}). Selection is of the top 20% males and females from 250 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. In addition to a deterministic model, results are presented for three stochastic models with different models for the polygenic component: the infinitesimal model, and finite locus models with 50 or 10 unlinked loci of equal effect but frequencies drawn from a uniform [0,1] distribution. Stochastic simulation results are the average of 500 replicate simulations.

in frequency of the QTL were very similar between genetic models (results not shown).

For the infinitesimal model, differences in response between MAS and phenotypic selection were very similar for the stochastic model and the deterministic model (Fig. 14.3). In contrast to the deterministic model, the stochastic model accommodates reductions in polygenic variance as a result of the Bulmer effect and inbreeding (Fig. 14.4). Under the stochastic model, however, changes in polygenic variance were similar for MAS and phenotypic selection and would, therefore, have limited impact on their contrast.

The finite locus model with 50 polygenes exhibited similar differences in cumulative response between MAS and phenotypic selection as the

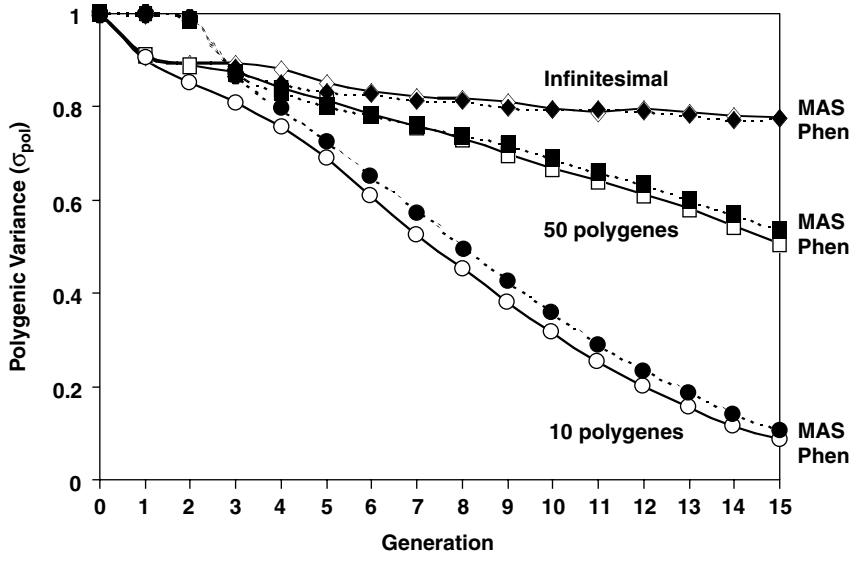

Fig. 14.4. Polygenic variance (relative to polygenic variance in generation 0) under standard MAS and phenotypic selection (Phen). Selection is of the top 20% males and females from 250 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. Results are presented for three models for the polygenic component: the infinitesimal model, and finite locus models with 50 or 10 unlinked loci with equal effect but frequencies drawn from a uniform [0,1] distribution. Results are the average of 500 replicate stochastic simulations.

infinitesimal model for the first 10 generations (Fig. 14.3). In subsequent generations, MAS regained some of the response it had lost under the finite locus model and differences with phenotypic selection decreased slightly. The recovery of lost response under MAS was greater for the model with 10 polygenes; the difference with phenotypic selection was reduced to 0.13 polygenic standard deviations ($=h_p\sigma'_p$) by generation 15. This behavior of the finite locus model is explained by the change in frequencies of polygenes over generations, as illustrated in Fig. 14.5. The average frequency of polygenes is initially lower for MAS than phenotypic selection. As frequencies move closer to 1, however, polygenic variance is depleted, and more rapidly so for phenotypic selection than for MAS (Fig. 14.4). As a result, polygenic response with MAS is able to catch up with polygenic response for phenotypic selection.

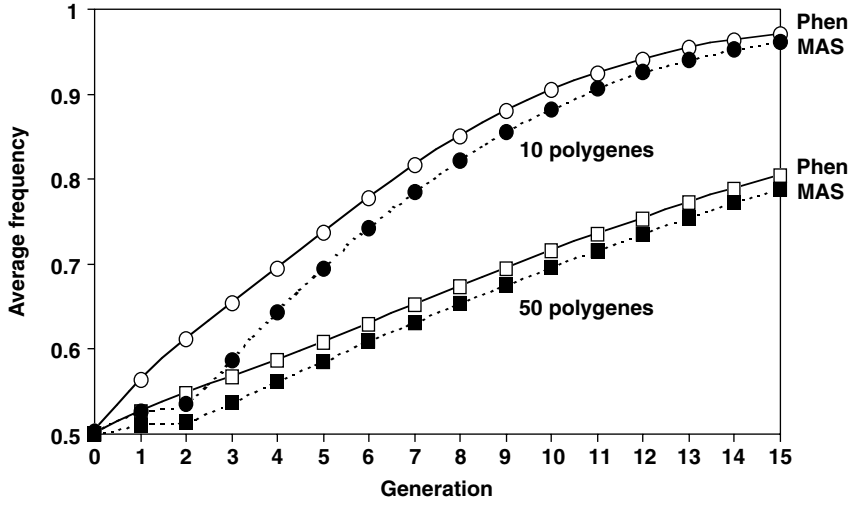

Fig. 14.5. Average frequency of polygenic loci under standard MAS and phenotypic selection (Phen). Selection is of the top 20% of males and females from 250 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. Polygenes are simulated based on 50 or 10 unlinked loci with equal effect but frequencies drawn from a uniform [0,1] distribution. Results are the average of 500 replicate stochastic simulations.

In a large population with negligible inbreeding, all genes that affect the trait will ultimately be fixed for their favorable allele under both MAS and phenotypic selection. Thus, ultimate response will be the same for both strategies. In populations of limited size, however, ultimate response will differ between strategies because of their impact on rates of fixation and loss of polygenes. Thus, ultimate differences in response between MAS and phenotypic selection will depend on the proportion of polygenes for which the favorable allele is lost. This is illustrated in Fig. 14.6, where for the example situation a slightly greater proportion of polygenes was lost with MAS than with phenotypic selection because of the reduced emphasis on polygenes in early generations. This difference was greater for the model with 10 polygenes than for the model with 50 polygenes. Thus, although MAS may recover some of the lost polygenic response, ultimate response will remain lower than for phenotypic selection because of the greater loss of favorable polygenes.

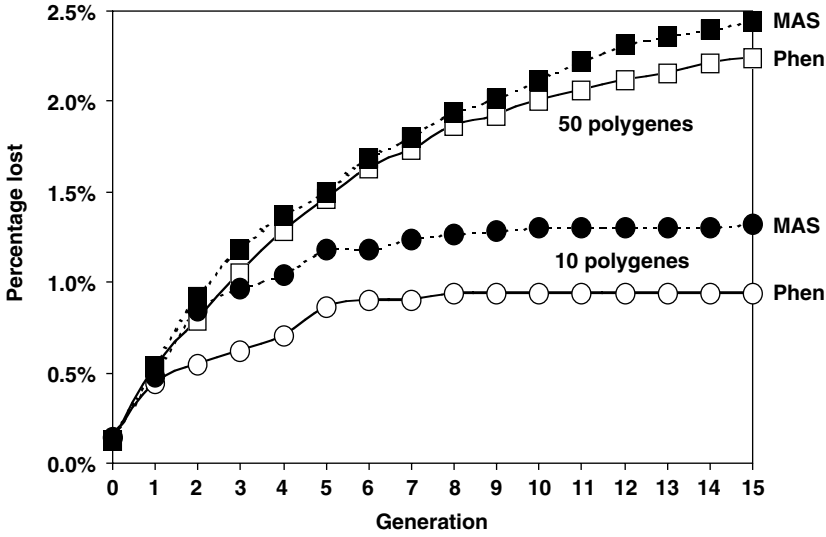

Fig. 14.6. The percentage of polygenic loci that are lost under standard MAS and phenotypic selection (Phen). Selection is of the top 20% of males and females from 250 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. Polygenes are simulated based on 50 or 10 unlinked loci with equal effect but frequencies drawn from a uniform [0,1] distribution. Results are the average of 500 replicate stochastic simulations.

III. OPTIMIZING MARKER-ASSISTED SELECTION

A. Optimal Selection on Single QTL

The previous results demonstrate that MAS strategies that maximize response over a single generation will not maximize response over more than one generation. The underlying reason is that selection not only changes the population mean but also population parameters such as gene frequencies and genetic variances. These changes in population parameters affect the amount of response that can be made in subsequent generations. Thus, strategies that maximize response over multiple generations must account for changes in population parameters that affect subsequent responses to selection.

In the case of MAS and under the infinitesimal genetic model with constant polygenic variance, the main population parameter that affects response to selection in subsequent generations is the frequency of the

QTL. Therefore, to develop strategies that maximize longer-term response to selection, Dekkers and van Arendonk (1998) used the deterministic model described previously to optimize weights in the following index of the genetic value for a single known QTL (g_{Qi}) and an EBV for polygenes:

$$I^{opt}_{i,Q,t} = b_{Q,t}g_{Qi} + \hat{g}_{pi}$$

Index weights $b_{Q,t}$ were allowed to differ by generation, sex, and QTL genotype. In reference to Fig. 14.1, changing weights on the QTL changes the means of the three distributions and, thereby, the proportions selected from each genotype.

Dekkers and van Arendonk (1998) used optimal control theory to derive the index weights that maximized cumulative response after T generations. Optimal control theory utilizes the unique structure of response to selection over generations, in that the optimal selection strategy for generation t depends only on population parameters in generation t, that is, polygenic means and QTL frequency, and not on the path that led to these parameters (Dekkers and van Arendonk 1998). Manfredi et al. (1998) solved a similar problem using a more general optimization method. Their method does not utilize the unique structure of selection over multiple generations and requires more computing time. The approach of Dekkers and van Arendonk (1998) was subsequently extended to multiple QTL by Chakraborty et al. (2002).

Fig. 14.7 illustrates the optimal weights assigned to the QTL in each generation for the example situation when the objective was to maximize cumulative response over 30 generations. Index weights were the same for males and females because selection intensities were the same for both. Weights differed by generation and QTL genotype. Except for the final generations, weights on the QTL were substantially lower than those used for standard MAS ($b_Q = 1$ for all generations) and phenotypic selection ($b_Q = h^2_p$ for all generations). Optimal weights were equal to 1 for the final generation because at that point the aim is to maximize response in the next generation, equivalent to standard MAS. In generation 29, the optimal weight on the unfavorable QTL genotype (bb) was extremely large.

Fig. 14.8 depicts the resulting changes in cumulative total and polygenic responses for optimal MAS as a deviation from responses to phenotypic selection. Frequencies of the QTL in each generation are illustrated in Fig. 14.8b. Optimal MAS led to a much more gradual and almost linear increase in frequency toward fixation at the end of the planning horizon. This is in contrast to standard MAS and phenotypic selection (Fig. 14.2b). As a result, cumulative response was lower for optimal MAS than for phenotypic selection (Fig. 14.8) for the first 23

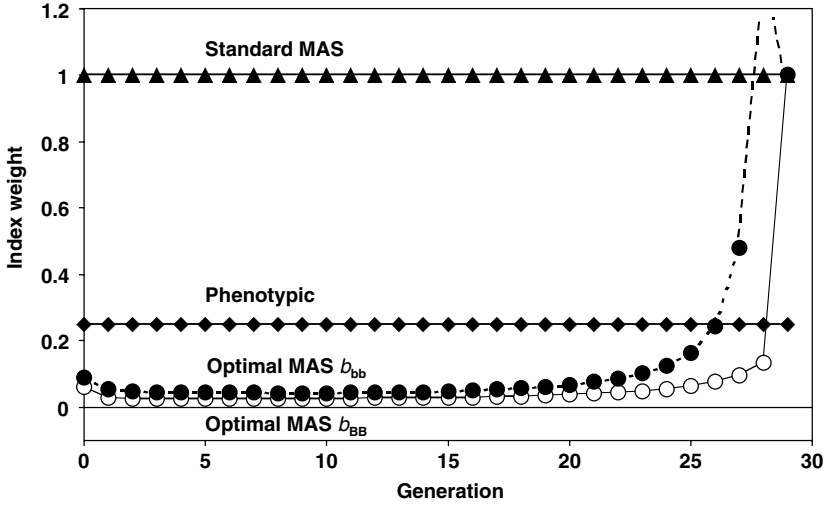

Fig. 14.7. Weights assigned to the QTL with standard MAS, phenotypic selection, and optimal MAS. Selection is of the top 20% of males and females for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. For optimal selection, b_{BB} and b_{bb} are the weights assigned to the favorable and unfavorable homozygotes, respectively. Weights on bb and BB are equal for standard MAS and phenotypic selection.

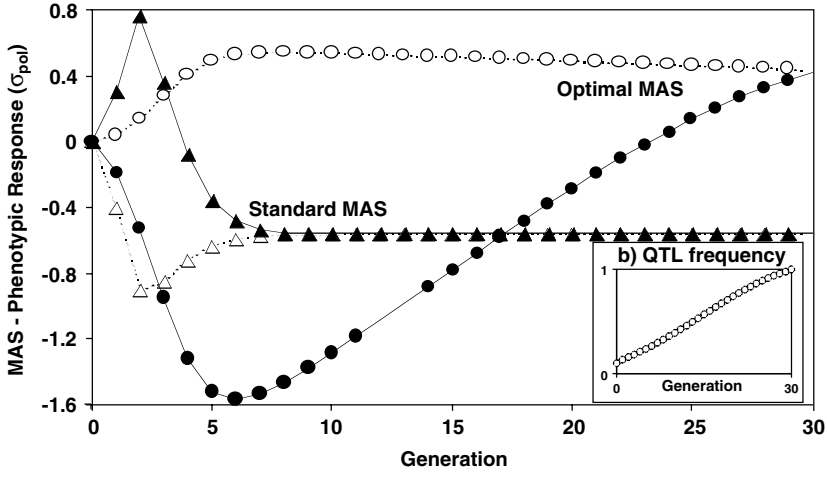

Fig. 14.8. Cumulative total (closed symbols) and polygenic (open symbols) responses to standard and optimal MAS in a population of infinite size, as a deviation from cumulative response for phenotypic selection (MAS response – phenotypic response, expressed in polygenic standard deviations, σ_{pol}), and b) frequency of the QTL for optimal MAS. Selection is of the top 20% of males and females for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. Results are based on the deterministic model.

generations. However, polygenic response was greater, which led to a 0.42 polygenic standard deviation greater cumulative response by the end of the planning horizon, at which time the QTL was also fixed under optimal MAS.

Realized selection intensities that were placed on the polygenes and the QTL in a generation, which were computed as response generated in that generation divided by the standard deviation of the selected component (= polygenic standard deviation for polygenic response and = $\sqrt{2p(1-p)}$ for the QTL, where p is the gene frequency), are given in Fig. 14.9. For optimal MAS, selection intensity placed on the polygenes was remarkably constant over generations, apart from the last generation. In contrast, selection pressure placed on polygenes was lower for both standard MAS and phenotypic selection prior to fixation of the QTL. Patterns for selection intensities placed on the QTL nearly mirrored those of intensity on polygenes for standard MAS and phenotypic selection but was again nearly constant for optimal MAS, apart from the first and last generation. The latter likely relates to the build-up of gametic phase disequilibrium between the QTL and polygenes.

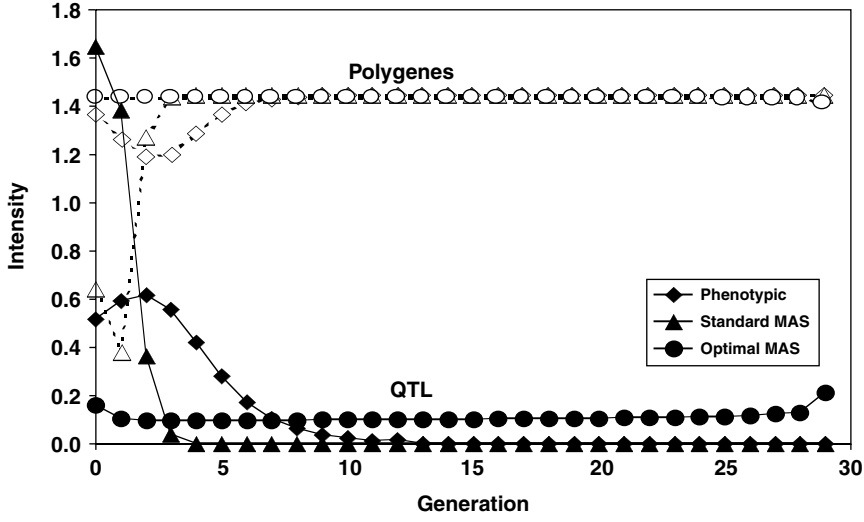

Fig. 14.9. Standardized selection response (intensity) per generation for the QTL (closed symbols) and polygenes (open symbols) under standard MAS, phenotypic selection, and optimal MAS. Selection is of the top 20% of males and females for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect $a = 1$ phenotypic standard deviations and frequency 0.1. Polygenic heritability is 0.25. Results are based on the deterministic model.

Results depicted in Fig. 14.9 demonstrate that, to maximize cumulative response over a planning horizon of T generations, selection emphasis on the QTL should be controlled in such a manner that the QTL is close to fixation in generation T, while equal selection emphasis is placed on polygenes across generations. Note that this optimal solution is by nature similar to that obtained by Finney (1958) for selection across multiple stages. He also found that to maximize cumulative response in a quantitative trait over multiple stages of selection, selection efforts should be divided equally across stages. Unequal selection results in lower cumulative response because of the non-linear relationship between proportion selected and selection intensity (Falconer and Mackay 1996). The additional complication in the present context is that the total selection emphasis that is applied to polygenes across all generations is not predetermined but must be balanced against placing sufficient emphasis on the QTL, such that the QTL frequency is moved to fixation in generation T. Results depicted in Fig. 14.9 show that this is achieved by maintaining a nearly constant standardized selection emphasis on the QTL over generations, apart from the first and last generations.

B. Optimal Selection on Multiple QTL

The deterministic optimization approach of Dekkers and van Arendonk (1998) was extended to multiple QTL by Chakraborty et al. (2002) and applied to simultaneous selection on two unlinked and linked QTL by Dekkers et al. (2002). To illustrate the properties of optimal selection on multiple QTL, the method of Chakraborty et al. (2002) was applied to selection on three unlinked additive QTL with parameters given in Table 14.1. Effects and starting parameters of the QTL were chosen to illustrate the impact of the effect and starting frequency of the QTL on optimal strategies.

Table 14.1. QTL effects and starting frequencies for the example with simultaneous selection on three biallelic and additive QTL.

QTL	QTL effect (a in phenotypic SD)	Initial frequency of the favorable allele
1	0.50	0.1
2	0.25	0.1
3	0.25	0.3

Cumulative responses to standard and optimal MAS, deviated from response to phenotypic selection (not shown), followed similar trends as depicted in Fig. 14.8 for the single QTL, except that the loss of cumulative response with standard MAS was nearly twice as large in generation 20 as it was for the single QTL. Extra response from optimal MAS over phenotypic selection was 0.26 polygenic SD in generation 20. The greater loss is related to the greater cumulative effect of the QTL.

Trends in frequencies of the QTL with standard and optimal MAS are illustrated in Fig. 14.10. With standard MAS, rates of fixation depended on both effect and starting frequency of the QTL and the QTL with the larger effect was moved to fixation more quickly, as expected. The same was observed for phenotypic selection (not shown), but rates of fixation were lower for all QTL. In contrast, with optimal MAS, frequencies increased nearly linearly to reach near fixation at the end of the planning horizon for all three QTL, regardless of the effect and the initial frequency of the QTL. The rate of increase in frequency was determined by the initial frequency of the QTL, not by its effect; magnitude of the QTL

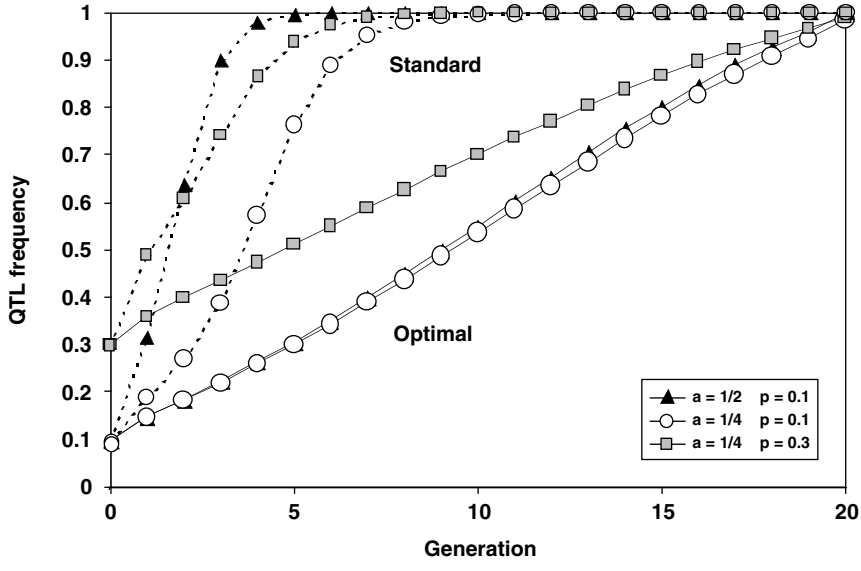

Fig. 14.10. Frequencies for three QTL under standard and optimal MAS. Selection is of the top 20% of males and females for a trait controlled by three unlinked biallelic additive QTL and polygenes. The QTL have with effects *a* (expressed in phenotypic standard deviations) and initial frequencies *p*. Polygenic heritability is 0.25. Results are based on the deterministic model.

had no impact on the selection emphasis that was placed on the QTL in the optimal strategy. This result is further illustrated in Fig. 14.11, which gives the realized selection intensities that were placed on the QTL and polygenes. Similar to optimal selection on a single QTL (Fig. 14.9), optimal selection pressure on polygenes was nearly constant over generations. Optimal selection intensities on the QTL depended only on initial frequency and were nearly constant over generations, apart from the first and last generation.

C. Optimizing Selection on QTL in Small Populations

The previous results were based on a deterministic model that assumes a large population (no inbreeding) and the infinitesimal genetic model with constant variance for polygenes. In that case, the optimal strategy to select on the QTL to maximize long-term response is to maintain equal selection emphasis on the polygenes over generations, such as to maximize polygenic response, and to increase the frequency of QTL to fixation at the end

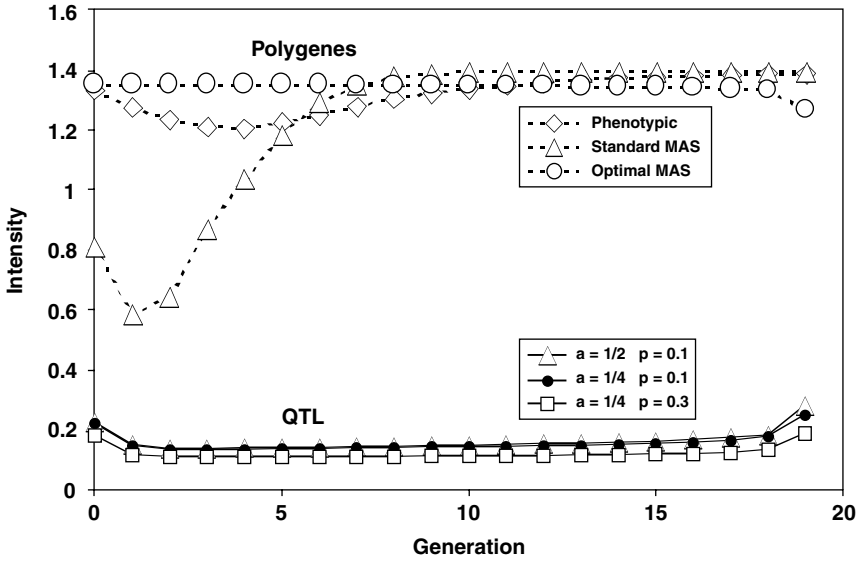

Fig. 14.11. Standardized selection response (intensity) per generation for polygenes under standard MAS, phenotypic selection, and optimal MAS, and for three QTL under optimal MAS. Selection is of the top 20% of males and females for a trait controlled by three unlinked biallelic additive QTL and polygenes. The QTL have effects *a* (expressed in phenotypic standard deviations) and initial frequencies *p*. Polygenic heritability is 0.25. Results are based on the deterministic model.

of the planning horizon in a nearly linear manner. For a very long planning horizon, this implies that very little emphasis should be placed on the QTL. Most long-term selection experiments are, however, based on populations of limited size, which provides additional complications for selection on known QTL. To investigate these issues, alternative selection strategies were simulated for 100 generations in a population of 100 individuals (50 males and 50 females) with selection of the best 10 males and females. Selection was for a trait affected by a single known additive QTL and polygenes that were simulated based on 50 unlinked loci of equal effect and frequencies sampled from a uniform distribution. The known QTL had a starting frequency equal to 0.3 and an effect $a = \frac{1}{2}\sigma_p'$. Polygenic heritability was 0.25. In addition to standard MAS and phenotypic selection, MAS strategies that more slowly increased the frequency of the QTL, similar to optimal MAS, were implemented. The latter were based on selection on the following index:

$$I_{i,Q,t} = k \frac{g_{Q_i}}{\sqrt{p_t(1 - p_t)}} + h_p^2 P_i'$$

where p_t is the frequency of the QTL in generation t. This index was designed to result in a constant selection pressure on the QTL by dividing the QTL effect, g_{Qi}, by the standard deviation of the QTL, such as to approximate optimal selection, with the overall emphasis on the QTL controlled by the constant k. Different values of the constant k were evaluated. A value of $k = 0.07$ resulted in the greatest cumulative response over 100 generations. All strategies were evaluated for replication 100 times.

Trends in QTL and average polygenic frequencies for phenotypic selection, standard MAS, and two implementations of "optimal" MAS, with $k = 0.02$ and $k = 0.07$, are shown in Fig. 14.12. By generation 80, alleles for the QTL and polygenes were fixed at 0 or 1 and no genetic variance remained. Differences in ultimate response to selection are, therefore, a result of the proportion of favorable alleles that were fixed at 1.

The strategy of MAS with $k = 0.07$ resulted in fixation of the QTL for all replicates prior to generation 40 (Fig. 14.12). The increase in the average QTL frequency was much slower for $k = 0.02$ and never reached 1. For this low selection emphasis on the QTL, the favorable allele was lost in over 20% of replicates, although the starting frequency was relatively high at 0.3. As a result, cumulative response after 100 generations was lower for $k = 0.02$ than for standard MAS, as illustrated in Fig. 14.13.

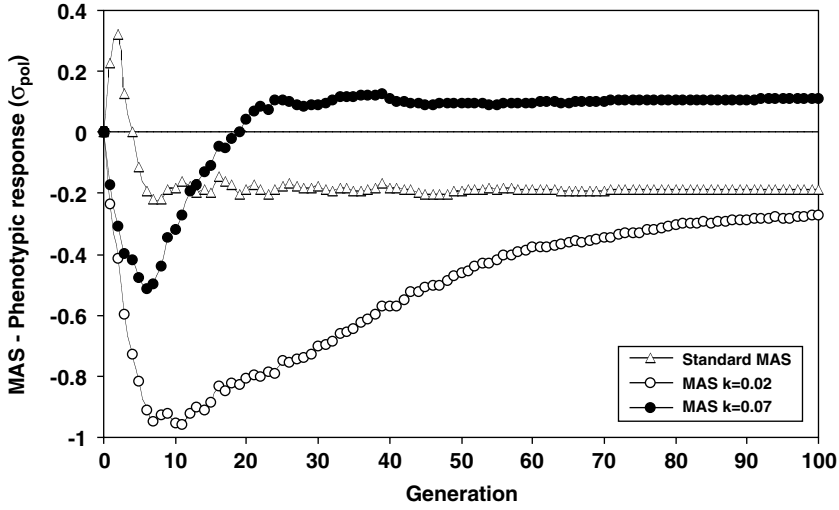

Fig. 14.12. Cumulative responses to standard MAS and to MAS with two levels of emphasis on the QTL (k = 0.02 and 0.07), as a deviation from cumulative response for phenotypic selection (MAS response – phenotypic response, expressed in polygenic standard deviations, σ_{pol}). Selection is of the top 20% of males and females from 50 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect a = 0.5 phenotypic standard deviations and frequency 0.3. Polygenic heritability is 0.25. Polygenes are simulated by 50 unlinked loci of equal effect but frequencies drawn from a uniform [0,1] distribution. Stochastic simulation results are the average of 100 replicate simulations.

Greatest long-term response was obtained with MAS with k = 0.07, although the difference with phenotypic selection was less than 0.2 polygenic standard deviations. The greater long-term response for MAS with k = 0.07, though small, was caused by a slightly lower loss of favorable polygenes (Fig. 14.14), 9.0% compared to 9.5% for phenotypic selection, and 10.2% for standard MAS. MAS with k = 0.02 lost 9.1% of polygenes, similar to MAS with k = 0.07.

D. Strategies to Control Drift During Marker-assisted Selection

Results in the previous section illustrate that in small populations random drift, and the associated chance of losing favorable alleles, is one of the main detriments to maximizing long-term response to selection. It is well known that random change in gene frequencies due to drift is

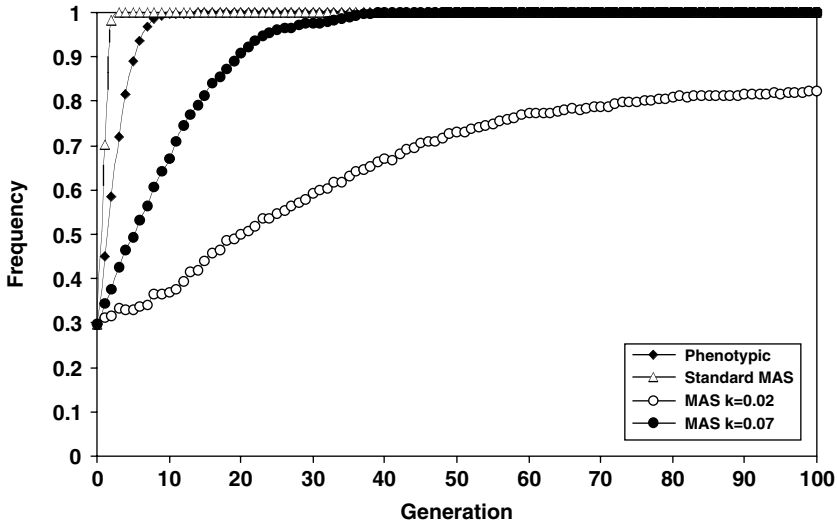

Fig. 14.13. Average frequency of the QTL under standard MAS and MAS with two levels of emphasis on the QTL (k = 0.02 and 0.07). Selection is of the top 20% of males and females from 50 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect a = 0.5 phenotypic standard deviations and frequency 0.3. Polygenic heritability is 0.25. Polygenes are simulated by 50 unlinked loci of equal effect but frequencies drawn from a uniform [0,1] distribution. Stochastic simulation results are the average of 100 replicate simulations.

an important determinant of long-term response with selection on phenotype (Walsh 2003) because it determines the probability of ultimate fixation versus loss of favorable alleles. With selection on phenotype, drift can be controlled by balancing the effects of inbreeding and selection intensity (Robertson 1960).

With index selection on a QTL, long-term response in polygenes can be controlled to some degree by decreasing selection emphasis on the QTL, thereby increasing selection pressure on polygenes. Reduced selection pressure on the QTL, which was found to be optimal in large populations (Fig. 14.7), however, increases the chance of loss of the QTL. Reduced emphasis on the QTL likely also results in greater selection effort that must be applied over generations to move the QTL to fixation, which reduces the selection emphasis that can be placed on the polygenes.

One opportunity that is not utilized with index selection on QTL information is the opportunity to directly control random drift at the gene level; unlike what is possible for polygenes, changes in frequency

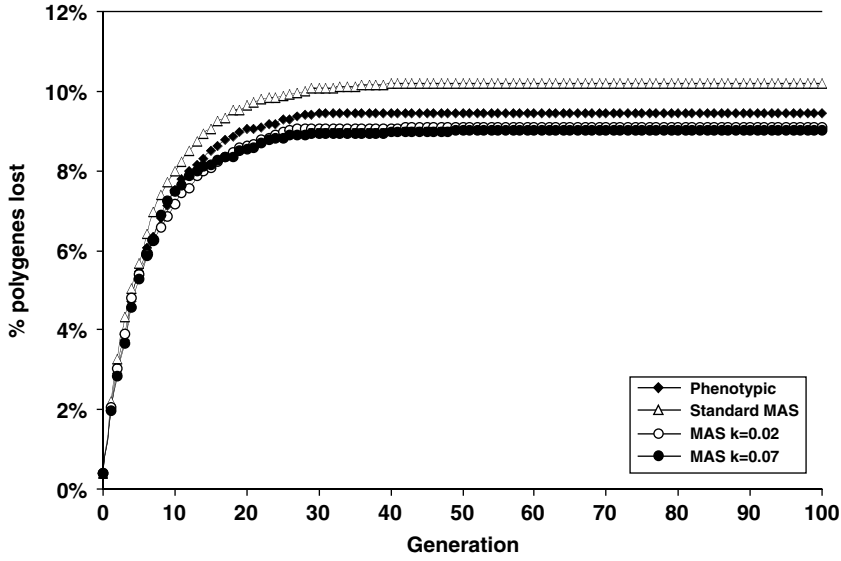

Fig. 14.14. Average percentage of polygenes lost under standard MAS and MAS with two levels of emphasis on the QTL (k = 0.02 and 0.07). Selection is of the top 20% of males and females from 50 individuals per sex for a trait controlled by a biallelic additive QTL and polygenes. The QTL has effect a = 0.5 phenotypic standard deviations and frequency 0.3. Polygenic heritability is 0.25. Polygenes are simulated by 50 unlinked loci of equal effect but frequencies drawn from a uniform [0,1] distribution. Stochastic simulation results are the average of 100 replicate simulations.

at the QTL can be controlled more precisely; if not directly, then indirectly through observed genotypes of linked genetic markers. One strategy to control random drift for QTL was suggested and evaluated by Hospital et al. (2000) in a program with exclusive selection on a molecular score for multiple QTL. This strategy, which they called complementation selection, involved truncation selection on the molecular score but with a minimum constraint on the number of selected parents that carry a given QTL. This strategy avoided loss of QTL and resulted in greater rates of fixation of favorable QTL alleles than simple truncation selection on molecular score.

An alternative that accomplishes the same objective but could lead to greater cumulative polygenic response would be a desired gains strategy, in which parents are selected to achieve the desired change in QTL frequency in the progeny generation, while maximizing polygenic response. Desired QTL frequencies could be based on the linear trends to fixation that are derived for optimal MAS. In addition to limiting the

risk of loss of favorable QTL alleles, controlling drift for the QTL may also maximize the emphasis that can be placed on polygenes. This will minimize the loss of favorable polygenes and maximize the probability of ultimate fixation of new favorable mutations.

In the above strategies, drift at polygenes could be further reduced by using optimal contribution selection (e.g., Villanueva et al. 1999), in which response to selection is maximized with a constraint on the rate of inbreeding. In addition, anonymous markers spread over the genome could be used to reduce drift at polygenes. Further research in these areas is warranted.

IV. CONCLUSIONS AND IMPLICATIONS

The main conclusion of the work presented here is that standard strategies for the use of molecular information in selection programs that maximize short-term response are expected to reduce long-term response to selection compared to selection on phenotype alone. However, strategies for MAS can be derived that optimize the use of molecular data to maximize long-term response. Such strategies are expected to result in greater long-term response than phenotypic selection, although the differences are expected to be limited for traits with moderate heritability and when phenotype is observed in all individuals prior to selection. The optimal use of molecular data to maximize long-term response involves controlling progress and drift for the QTL, such that emphasis on polygenes is maximized over generations and the probability of loss of polygenes is minimized. For the QTL, this implies increasing the frequency of the QTL at a nearly constant rate to fixation at the end of the planning horizon, regardless of the effect of the QTL. Depending on length of planning horizon, this means placing considerably less emphasis on the QTL than with standard MAS and even compared to the implicit emphasis that is placed on the QTL with phenotypic selection. The result is lower short-term response than with standard MAS, because the short-term advantage of rapidly increasing the frequency of the QTL is not capitalized on, but greater long-term response because cumulative response for polygenes is maximized.

The interest of most commercial breeding programs is in the short term, for obvious economic reasons. Thus, strategies for MAS that maximize long-term response to selection should not be recommended for use in commercial breeding programs. Nevertheless, it is important for breeders to realize that short-term breeding decisions have important implications for responses that can be obtained in the future. This has

been realized for several decades when it comes to balancing selection intensity and inbreeding in recurrent selection programs; short-term response is increased by increasing selection intensity but this will reduce longer-term response to selection because of the ensuing rates of inbreeding and associated loss of variance. The optimal balance between selection intensity and inbreeding to maximize long-term response under phenotypic selection, as derived by Robertson (1960), is to select half of the population, which is much more than required for rapid short-term response. Similar considerations hold for the use of molecular data in recurrent selection through MAS. An additional consideration with MAS, however, is the greater genotyping cost that is incurred when fixation of the QTL is delayed. These costs would put further emphasis in commercial programs on maximizing short-term gain and fixing the QTL rapidly.

One solution to the antagonism between short- and long-term response from a commercial perspective is to maximize a combination of short- and long-term response. Such strategies were considered by Dekkers and Chakraborty (2001), who derived and evaluated strategies for selection on a QTL that maximized a discounted sum of responses in each generation of a planning horizon. In this criterion, early responses receive more weight than later responses, depending on the discount rate chosen.

Results presented herein were based on several simplifying assumptions that need further consideration. One was that the QTL could be genotyped directly. In practice, often only markers linked to QTL are available. This will require additional emphasis on the markers in early generations before the disequilibrium between the marker and the QTL is lost. Also, effects of the QTL must be estimated and resulting estimates tend to be biased upward (Beavis 1994; Melchinger et al. 1998; Bost et al. 2001). Although QTL effects have no impact on the selection strategy if the aim is to maximize long-term response (Fig. 14.11), estimates must be used if the aim is to maximize short-term response.

In the present study, QTL and polygenes were assumed to be unlinked. If polygenes are linked to QTL, which is likely, standard MAS can increase the loss of favorable linked polygenes through linkage drag. To reduce this possibility, early selection on the QTL must focus on selection of recombinants in regions adjacent to the QTL (Hospital 2001). The QTL can also be linked to each other. This was investigated by Dekkers et al. (2002), who showed that selection must be for parents that are in a favorable coupling phase, that is, AB/ab vs. Ab/aB. Selection of parents with tightly linked QTL in repulsion phase will result in lost selection emphasis on polygenes, because such chromosomes must be selected against eventually, unless there is sufficient recombination.

An additive genetic model was assumed throughout. The impact of dominance at the QTL on standard and optimal MAS on short- and medium-term responses was investigated by Dekkers (1999) and Dekkers and Chakraborty (2001). Dominance at the QTL makes standard MAS suboptimal even over one generation (Dekkers 1999). Dominance is not expected to affect strategies for optimal MAS that maximize long-term response unless the QTL shows overdominance. With overdominance, the QTL should not be fixed in the long term since heterozygotes have greatest value. Similar considerations hold for the impact of epistatic interactions between QTL and of QTL with polygenes on strategies for MAS that maximize long-term response; if the QTL genotype that maximizes phenotype is composed of homozygotes at all QTL, the optimal strategy for long-term response will be to fix all QTL at rates described herein. Known epistatic interactions between QTL can be accommodated in the methods of Chakraborty et al. (2002) to derive optimal MAS to maximize short- and medium-term responses.

LITERATURE CITED

Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals. Nature Rev. Genetics 2:130–138.

Beavis, W. D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. p. 250–266. In: 49th Ann. Corn and Sorghum Res. Conf. ASTA, Washington, DC.

Bost, B., D. de Vienne, F. Hospital, L. Moreau, and C. Dillmann. 2001. Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. Genetics 157:1773–1787.

Bulmer, M. G. 1980. The mathematical theory of quantitative genetics. Clarendon Press, Oxford.

Chakraborty, R., L. Moreau, and J. C. M. Dekkers. 2002. A method to optimize selection on multiple identified quantitative trait loci. Genet. Sel. Evol. 34:145–170.

Dekkers, J. C. M. 1999. Breeding values for identified quantitative trait loci under selection. Genet. Sel. Evol. 31:421–436.

Dekkers J. C. M., and R. Chakraborty. 2001. Potential gain from optimizing multi-generation selection on an identified quantitative trait locus. J. Animal Sci. 79:2975–2990.

Dekkers, J. C. M., R. Chakraborty, and L. Moreau. 2002. Optimal selection on two quantitative trait loci with linkage. Genet. Sel. Evol. 34:171–192.

Dekkers, J. C. M., and F. Hospital. 2002. Utilization of molecular genetics in genetic improvement of plants and animals. Nature Rev. Genetics 3:22–32.

Dekkers, J. C. M., and J. A. M. van Arendonk. 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. Genet. Res. 71:257–275.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Harlow, UK.

Fernando, R. L., and M. Grossman. 1989. Marker-assisted selection using best linear unbiased prediction. Genet. Sel. Evol. 21:467–477.

Finney, D. J. 1958. Statistical problems of plant selection. Bul. Int. Statist. Inst. 36:242–268.

Gibson, J. P. 1994. Short-term gain at the expense of long-term response with selection of identified loci. Proc. 5th World Cong. Genet. Appl. Livest. Prod. Univ. Guelph, Canada, 21:201–204.

Gimelfarb, A., and R. Lande. 1994. Simulation of marker assisted selection in hybrid populations. Genet. Res. 63:39–47.

Hospital, F. 2001. Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. Genetics 158:1363–1379.

Hospital F., I. Goldringer, and S. Openshaw. 2000. Efficient marker-based recurrent selection for multiple quantitative trait loci. Genet. Res. 75:357–368.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756.

Larzul, C., E. Manfredi, and J. M. Elsen. 1997. Potential gain from including major gene information in breeding value estimation. Genet. Sel. Evol. 29:161–184.

Manfredi E., M. Barbieri, F. Fournet, and J. M. Elsen. 1998. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. Genet. Sel. Evol. 30:127–148.

Mauricio, R. 2001. Mapping quantitative loci in plants: uses and caveats for evolutionary biology. Nature Rev. Genetics 2:370–381.

Melchinger, A. E., H. F. Utz, and C. C. Schon. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Meuwissen, T. H. E., and M. E. Goddard. 1996. The use of marker haplotypes in animal breeding schemes. Genet. Sel. Evol. 28:161–176.

Pong-Wong, R., and J. A. Woolliams. 1998. Response to mass selection when an identified major gene is segregating. Genet. Sel. Evol. 30:313–337.

Robertson, A. 1960. A theory of limits in artificial selection. Proc. R. Soc. Lond. B 153:234–249.

Villanueva, B., R. Pong-Wong, B. Grundy, and J. A. Woolliams. 1999. Potential benefit from using an identified major gene and BLUP estimated breeding values in selection programmes. Genet. Sel. Evol. 31:115–133.

Walsh, B. 2003. Selection theory and limits. Long-term Selection Symposium, Urbana, IL (dates TK)

Zhang, W., and C. Smith. 1992. Computer simulation of markers-assisted selection utilizing linkage disequilibrium. Theor. Appl. Genet. 83:813–820.