Genet. Sel. Evol. 31 (1999) 421–436

Original article

Breeding values for identified quantitative trait loci under selection

Jack C.M. Dekkers

Department of Animal Science, 225C Kildee Hall, Iowa State University, Ames, IA, 50011, USA

(Received 26 March 1999; accepted 6 July 1999)

Abstract – The use of an identified quantitative trait locus (QTL) in selection requires the integration of breeding values (BV) for the known QTL with estimates of polygenic BV. For a QTL with two alleles, BV for the QTL are traditionally based on the allele substitution effect, $\alpha = a + d(q - p)$, where a and d are additive and dominance effects, and p and q are gene frequencies in the current generation. It is shown here that to maximize single generation response, BV for a QTL with dominance must be derived based on gene frequencies among selected mates rather than frequencies in the current (unselected) generation. Because selection affects gene frequencies that in turn affect optimal BV for the QTL, gene substitution effects must be derived numerically. Response from selection on optimized versus standard BV for the QTL was evaluated for a range of parameters. Benefits of optimal selection were greatest for intermediate gene frequency and increased with a magnitude of additive and dominance effects up to 9 %. Extra response was negligible for gene frequencies less than 0.05 or greater than 0.85. In conclusion, strategies for marker-assisted selection that aim to maximize short-term response must account for the effects of dominance and changes in gene frequency at the QTL on performance of future progeny. © Inra/Elsevier, Paris

marker-assisted selection / dominance / breeding values / quantitative trait loci

Résumé – Valeurs génétiques pour des loci quantitatifs identifiés en situation de sélection. L'utilisation d'un locus quantitatif (QTL) identifié en sélection nécessite l'intégration des valeurs génétiques (BV) pour le QTL connu avec les estimées des BV polygéniques. Pour un QTL avec deux allèles, les BV à un QTL sont traditionnellement basées sur l'effet de substitution allélique, $\alpha = a + d(q - p)$, où a et d sont les effets additifs et de dominance et où p et q sont les fréquences géniques à la génération présente. On montre ici que pour maximiser la réponse à une seule génération de sélection, les BV pour un QTL avec dominance doivent être calculées à partir des fréquences géniques parmi les conjoints sélectionnés plutôt que des fréquences dans la génération présente non sélectionnée. Parce que la sélection

E-mail: jdekkers@iastate.edu

affecte les fréquences géniques qui à leur tour affectent les BV optimales pour le QTL, les effets de substitution de gènes doivent être calculés numériquement. La réponse à la sélection sur la valeur génétique optimisée ou classique pour le QTL a été évaluée pour une série de paramètres. Les bénéfices de la sélection optimale ont été plus importants pour les fréquences de gène intermédiaires et ont augmenté jusqu'à 9 % avec l'importance des effets additifs et de dominance. La réponse supplémentaire a été négligeable pour les fréquences géniques inférieures à 0,05 ou supérieures à 0,85. En conclusion, les stratégies de sélection assistée par marqueurs qui maximisent la réponse à court terme doivent tenir compte des effets de dominance et des changements de fréquence génique au QTL sur la performance de la descendance future.

sélection assistée par marqueurs / dominance / valeur génétique / locus quantitatif

1. INTRODUCTION

Permanent genetic improvement for quantitative traits is created by selection on the additive effects of genes that affect the trait of interest. Additive effects are termed breeding values and form the basis for genetic improvement programs in livestock and plants. An individual's breeding value is defined as the expected performance of progeny under random mating [4]. Selection can be made on estimates of the collective additive effects of genes on the trait without knowledge of the genes involved. Such estimated breeding values (EBV) can be derived based on phenotypic records of the individual and its relatives. To date, most programs for improvement of additive genetic merit in livestock and plants have relied on selection based on EBV derived from phenotypic records. Increasingly, however, information is becoming available on the effects of individual genes that affect quantitative traits, so-called quantitative trait loci (QTL). Information on QTL can be combined with EBV derived from phenotypic records to improve rates of genetic improvement.

Use of information from identified QTL (major genes) in selection for quantitative traits was first described by Neimann-Sorensen and Robertson [13]. They developed procedures to weight information from an identified QTL with phenotypic information using selection index procedures [8], based on the amount of genetic variance explained by the QTL. Smith [15] and Smith and Webb [16] extended these procedures and compared the rates of response from one generation of selection on this index to the response from selection on phenotypic information alone. Lande and Thompson [10] derived selection criteria combining information from genetic markers linked to QTL with phenotypic information, using the selection index theory. Marker information was combined into a marker score, which was equal to the sum of the average effects associated with markers. Average effects were defined as allele substitution effects and derived as partial coefficients of regression of phenotype on number of marker alleles [10]. Soller [17] considered the discrete nature of effects at an identified QTL in predicting response to selection. Selection was on an index of the breeding value for the QTL, which was assumed to be known without error, and an EBV for polygenic effects. Pong-Wong and Woolliams [14] showed that the discrete index used by Soller [17] is equivalent to the indexes of Smith [15] and Lande and Thompson [10] when QTL effects are known without error.

The indexes described by the above authors were designed to maximize the average genetic level of progeny when mated to a random group of unselected parents. In particular, the effects of identified QTL or markers used in these indexes were derived based on their average effects in an unselected population. In practical breeding programs, however, selection takes place in both sexes, and selected parents are mated to a selected rather than an unselected group of mates. This may change the average effect of alleles for genes that express dominance.

The impact of selection of mates on the breeding value for identified QTL was recognized by Larzul et al. [11], who developed a deterministic model for selection on a combination of an identified QTL and polygenes in a breeding program with overlapping generations. Breeding values and their estimates were obtained in an iterative manner within the context of the defined selection program. Larzul et al. [11], however, did not consider the nature of optimal breeding values for identified QTL, nor did they investigate the impact of the use of optimal versus standard breeding values for the identified QTL on selection response.

The objectives of this paper were, therefore, to derive breeding values for identified QTL that maximize the response to single generation selection and to evaluate the advantage of selection based on optimum breeding values over selection based on conventional breeding values for single genes. A single identified QTL with known effects is considered for simplicity, but implications for selection on marked QTL or when QTL effects are not known without error are discussed. The objectives of this paper are important relative to the use of information of individual genes in genetic improvement programs.

2. METHODS

2.1. Notation

Consider generation 0 of an unselected population of infinite size with discrete generations and in gametic phase equilibrium [1]. The population is recorded for a quantitative trait that is affected by an identified QTL and unlinked polygenes. All individuals are genotyped for the QTL prior to their age of selection. The QTL has two alleles, B and b, with frequencies p_0 and q_0. Following Falconer and MacKay [4], genotypic values for the QTL are a, d and $-a$ for individuals with genotypes G_i equal to BB, Bb and bb, respectively. Parameters and notation for the identified QTL are summarized in *table I*. Effects and frequencies of alleles at the QTL are assumed to be known without error.

Polygenic effects for the quantitative trait conform to the infinitesimal genetic model [4]. After accounting for effects at the identified QTL, the phenotypic standard deviation of the trait is σ_p and heritability is h^2. Breeding values for polygenic effects are estimated with accuracy r_m for males and r_f for females, resulting in a standard deviation of estimated breeding values for polygenic effects equal to $\sigma_m = r_m h \sigma_p$ for males and $\sigma_f = r_f h \sigma_p$ for females. With polygenic breeding values estimated based on own performance, $r_m = r_f = h$. The results derived here, however, apply to estimates of the polygenic breeding values derived based on selection index procedures, using

Table I. Frequencies, genotypic values and breeding values at a single locus with two alleles B and b with frequencies p_0 and (q_0), following Falconer and MacKay [4].

	BB	Bb	bb
Frequency	p_0^2	$2p_0q_0$	q_0^2
Genotypic value	a	d	$-a$
Breeding value	$2q_0\alpha_0$	$(q_0 - p_0)\alpha_0$	$-2p_0\alpha_0$
Adjusted breeding value*	$+\alpha_0$	0	$-\alpha_0$

* Adjusted breeding values are deviated from the breeding value of heterozygotes. α_0 is the average effect of gene substitution and is equal to $a + d(q_0 - p_0)$.

information from relatives, or based on the best linear unbiased prediction methods, with a model that includes a QTL genotype as a fixed effect (e.g. [9]).

Consider the selection of a fraction Q_s of males and Q_d of females to produce the next generation (generation 1). Mating of selected parents is at random. Selection is by truncation on an EBV that combines the breeding value for the identified QTL with an estimate of the polygenic breeding value:

$$\widehat{A}_{ijk} = g_{ij} + \widehat{u}_{ijk}, \tag{1}$$

where $\widehat{A}_{ijk}$ is the total EBV for animal k of sex j (male or female) and QTL genotype i (BB, Bb or bb), g_{ij} is the breeding value for the QTL for individuals with QTL genotype i of sex j, as a deviation from the QTL breeding value for individuals with genotype Bb ($g_{Bb,j} = 0$), and $\widehat{u}_{ijk}$ is an estimate of the polygenic breeding value for animal ijk. Following Falconer and MacKay [4], breeding values for the QTL for individuals with genotypes BB, Bb and bb are equal to $+2q_0\alpha_0$, $(q_0 - p_0)\alpha_0$, and $-2p_0\alpha_0$, where α_0 is defined as the average allele substitution effect and is equal to $\alpha_0 = a + (q_0 - p_0)d$. When selection is within a generation, QTL breeding values can for simplicity be deviated from the breeding value of the heterozygote without changing the ranking of individuals by subtracting $(q_0 - p_0)\alpha_0$. This results in adjusted QTL breeding values g_{ij} equal to $+\alpha_0$, 0 and $-\alpha_0$ (see *table I*).

2.2. Optimal QTL breeding values

Under random mating to selected mates, the EBV of an individual that is expected to maximize response from the current to the next generation can be derived as two times the expected mean of progeny conditional on the information available. Consider an individual ijk with QTL genotype G_i and polygenic EBV equal to $\widehat{u}_{ijk}$, which is mated at random to a group of mates with QTL gene frequencies p_m and q_m and average polygenic EBV equal to $\overline{u}_m$. Let p_i and q_i denote the frequency of gametes carrying the B and b allele: p_i equals 1, 1/2 and 0 for G_i equal to BB, Bb and bb, and $q_i = 1 - p_i$. Under random mating to selected mates, B and b gametes are combined at random to B and b gametes with frequencies p_m and q_m. Again taken as a deviation from the average EBV of Bb individuals, the EBV of an individual with QTL

genotype G_i can be derived as:

$$\begin{aligned}\widehat{A}_{ijk} &= 2E(P_{\text{prog}}|G_i,\ p_m, q_m, \widehat{u}_{ijk}, \overline{u}_m) - 2E(P_{\text{prog}}|Bb, p_m, q_m, \overline{u}_m) \\ &= 2[p_i p_m a + (p_i q_m + q_i p_m)d - q_i q_m a + 1/2\overline{u}_m + 1/2\widehat{u}_{ijk}] \\ &\quad - 2[1/2 p_m a + 1/2d - 1/2 q_m a + 1/2\overline{u}_m] \end{aligned} \tag{2}$$

Using the fact that $p_i + q_i = 1$ and $p_m + q_m = 1$, the latter equation can be simplified to:

$$\begin{aligned}\widehat{A}_{ijk} &= [(2p_i - 1)p_m + (1 - 2q_i)q_m]a + [2p_i q_m + 2q_i p_m^{-1}]d + \widehat{u}_{ijk} \\ &= 2(p_i - 1/2)[a + (q_m - p_m)d] + \widehat{u}_{ijk} \end{aligned} \tag{3}$$

Setting $\alpha_m = a + (q_m - p_m)d$ results in:

$$\widehat{A}_{ijk} = 2(p_i^{-1/2})\alpha_m + \widehat{u}_{ijk} \tag{4}$$

Resulting QTL breeding values are equal to $+\alpha_m$, 0 and $-\alpha_m$ for individuals with genotypes BB, Bb and bb. Note that this result is consistent with the quantitative genetic theory [4, 17], except that the gene substitution effect α_m is based on gene frequencies among selected mates rather than frequencies among all selection candidates.

Letting p_s and q_s be the frequencies of B and b among selected males and p_d and q_d the frequencies among selected females, optimal breeding values for the QTL become equal to $+\alpha_s$, 0 and $-\alpha_s$ for sires and equal to $+\alpha_d$, 0 and $-\alpha_d$ for dams, with:

$$\alpha_s = a + (q_d - p_d)d \tag{5}$$

and

$$\alpha_d = a + (q_s - p_s)d \tag{6}$$

2.3. Numerical procedures for derivation of optimal QTL breeding values

The problem with the implementation of the procedures described above for selection on the identified QTL is that optimal breeding values for the QTL in index (4) depend on the gene frequency of the QTL among mates, which in turn depends on the selection that takes place among mates and, therefore, on the index used for selection. This means that optimum breeding values cannot be derived analytically, but that iterative procedures are required. These procedures, which are derived below, involve the prediction of gene frequencies among selected sires and dams for given values of α_s and α_d, followed by updating α_s and α_d in an iterative manner based on the new frequencies among selected sires and dams.

2.3.1. *Deterministic model for selection on given QTL breeding values*

For each sex, truncation selection on index $\widehat{A}_{ijk} = 2(q_i - 1/2)\alpha_m + \widehat{u}_{ijk}$ involves selection across three Normal distributions that correspond to individuals with QTL genotypes BB, Bb and bb, as illustrated in *figure 1.* Distributions have means equal to $+\alpha_s$, 0 and $-\alpha_s$ for sires and equal to $+\alpha_d$, 0 and $-\alpha_d$ for dams. The standard deviation of the three distributions is equal to σ_m for males and σ_f for females. The frequency of each distribution is determined by the frequency of QTL genotypes among selection candidates, which under random mating is equal to p_0^2, $2p_0q_0$ and q_0^2 in generation 0.

For a given set of frequencies, means (based on α_j) and standard deviations of the three distributions, a unique truncation point c_j exists across the three distributions for sex j that results in the correct selected fraction (Q_s for males and Q_d for females). Let f_{ij} and x_{ij} be the selected fraction and standardized truncation point, respectively, for the distribution of EBV for individuals with QTL genotype i of sex j. The unique truncation point on the EBV scale, c_j, relates to the standardized truncation points x_{ij} based on:

$$c_j = \mu_{ij} + x_{ij}\sigma_j \tag{7}$$

where μ_{ij} is equal to $+\alpha_j$, 0 and $-\alpha_j$ for genotypes BB, Bb and bb. Also, the following relationships must exist between the standardized truncation points x_{ij}:

$$x_{BB,j} - x_{Bb,j} = x_{Bb,j} - x_{bb,j} = \alpha_j/\sigma_j \tag{8}$$

In addition, the f_{ij} fractions selected from distribution ij, which are equal to $1 - \Phi(x_{ij})$, where Φ is the cumulative distribution function for a standard normal distribution, must satisfy a constraint on the overall fraction selection:

$$f_{BB,j}p_0^2 + f_{Bb,j}2p_0q_0 + f_{bb,j}q_0^2 = Q_j \tag{9}$$

Equations (7)–(9) uniquely define the truncation point c_j. Even for given distribution parameters, an analytical solution does not exist but c_j must be solved iteratively. Iteration can be based on a Newton method algorithm, as developed by Ducrocq and Quaas [3], or on a bisection method as suggested by Gibson [6] and given in Appendix I. Once the unique truncation point c_j has been obtained, the QTL frequency among selection candidates (p_s and p_d) can be derived from

$$p_j = \frac{1}{Q_j}\{f_{1j}p_0^2 + f_{2j}p_0(1 - p_0)\} \tag{10}$$

With random mating of selected parents, the average genetic value of progeny can be derived based on

$$\overline{A}_1 = (p_s + p_d - 1)a + (p_s + p_d - 2p_sp_d)d + \overline{u}_1 \tag{11}$$

where $\overline{u}_1$ is the average polygenic value of progeny. This value $\overline{u}_1$ can be predicted using standard methods of predicting response to selection pooled across

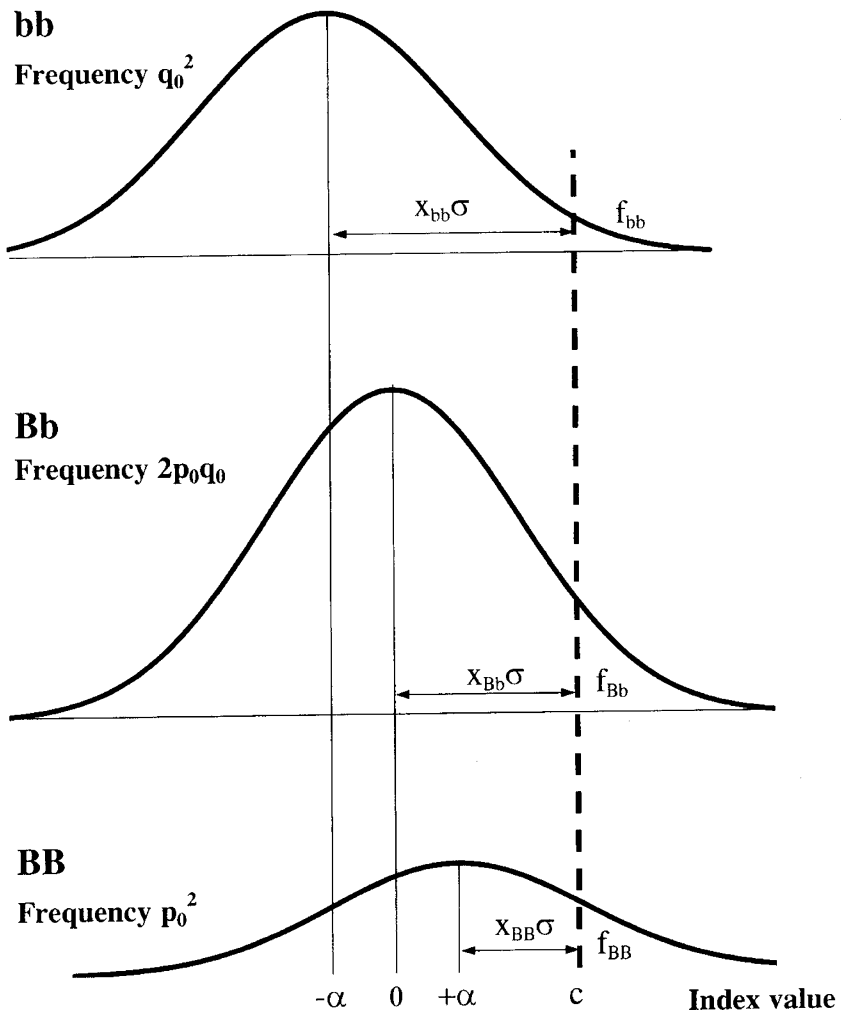

Figure 1. The process of selection on an index that combines information from a QTL with estimates of polygenic breeding values. The QTL has two alleles (B and b) with frequencies p_0 and q_0 and allele substitution effect α. Estimates of polygenic breeding values have a standard deviation equal to σ. Selection on the index is by truncation across three Normal distributions at truncation point c on the index scale and, for QTL genotype j, at standardized truncation points x_j and selected fractions f_j.

QTL genotypes and sexes as:

$$\overline{u}_1 = \overline{u}_0 + \frac{\sigma_s}{2Q_s}\{p_0^2 f_{1s} i_{1s} + 2p_0(1-p_0) f_{2s} i_{2s} + (1-p_0)^2 f_{3s} i_{3s}\}$$
$$+ \frac{\sigma_d}{2Q_d}\{p_0^2 f_{1d} i_{1d} + 2p_0(1-p_0) f_{2d} i_{2d} + (1-p_0)^2 f_{3d} i_{3d}\} \qquad (12)$$

where i_{ij} is the selection intensity for genotype i from sex j.

2.3.2. *Iterative procedure for deriving optimal QTL breeding values*

Iterative procedures for finding the unique truncation points for given allele substitution effects must be incorporated within an iterative procedure for finding the optimal QTL substitution effects α_s and α_d. The following procedure can be used:

1) Set $\alpha_s = \alpha_d = a + (q_0 - p_0)d$.

2) Set $x_{BB,s} - x_{Bb,s} = x_{Bb,s} - x_{bb,s} = \alpha_s/\sigma_s$ and $x_{BB,d} - x_{Bb,d} = x_{Bb,d} - x_{bb,d} = \alpha_d/\sigma_d$.

3) Find the unique truncation points c_s and c_d and fractions selected, f_{ij}, based on the procedures described in section 2.3.1.

4) Compute the frequency of QTL alleles among selected parents p_s and p_d, based on equation (10).

5) Using the new solutions for p_s and p_d, compute new values for α_s and α_d as: $\alpha_s = a + (q_d - p_d)d$ and $\alpha_d = a + (q_s - p_s)d$. A multiplicative relaxation factor may be required here, reducing changes in α_s and α_d from one iteration to another, to allow convergence.

6) Repeat steps 2 through 5 until α_s and α_d converge to stable solutions.

Once optimal solutions have been obtained, the expected genetic level among progeny can be determined based on equations (11) and (12). Note that the starting values for this iterative procedure, which are set in step 1, provide results for classic selection with a known QTL.

2.4. Optimal QTL breeding values with gametic phase disequilibrium

In section 2.3, the parental generation was assumed to be in gametic phase equilibrium. When gametic phase disequilibrium is present in the parental population as a result of prior selection, average polygenic values will differ by QTL genotype; with truncation selection, individuals with the favorable QTL genotype tend to have lower polygenic values. This disequilibrium must be incorporated in selection decisions. Let $\overline{u}_{ij}$ be the average polygenic breeding value for QTL genotype i for sex j. Under random mating of selected parents, average polygenic values will be equal for male and female progeny and equidistant between the three progeny genotypes, and hence let $\overline{u}_{BB,j} = \overline{u}_{Bb,j} - \overline{u}_{Bb,j} - \overline{u}_{bb,j} = \delta$. Assuming δ can be estimated with sufficient accuracy, gametic phase disequilibrium between the QTL and polygenes can be accounted for in the selection index as follows (e.g. [14])

$$\widehat{A}_{ijk} = 2(p_i - 1/2)[a + (q_m - p_m)d + \delta] + \widehat{u}_{ijk} \tag{13}$$

where $\widehat{u}_{ijk}$ is now the individual's polygenic EBV as a deviation from the average polygenic breeding value for individuals of QTL genotype i and sex j. Based on this, optimal QTL allele substitution effects can be derived as before but with the effect of gametic phase disequilibrium included in the allele substitution effect as:

$$\alpha_s = a + (q_d - p_d)d + \delta \tag{14}$$

$$\alpha_d = a + (q_s - p_s)d + \delta \tag{15}$$

Note that because δ is negative, a gametic phase disequilibrium will reduce the average allele substitution effect associated with the QTL.

3. RESULTS

3.1. One generation response

Methods for the optimization of single generation response were applied and the responses were compared to selection on an index in which breeding values for the QTL were derived based on frequency in the parental generation ($\alpha = a + (1 - 2p_0)d$). These two strategies will be referred to as optimal and standard gene-assisted selection (GAS), respectively.

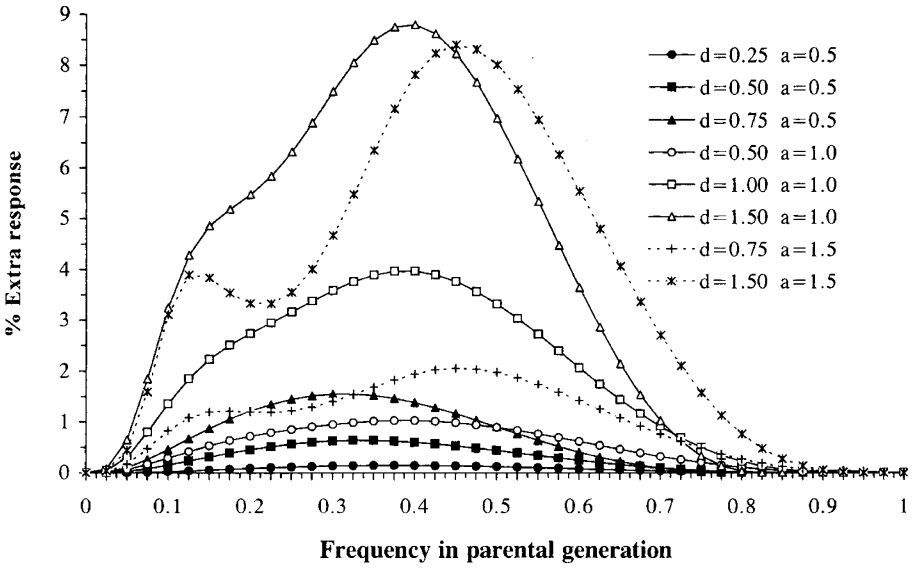

Figure 2. Extra response (%) from single generation selection on an index that combines breeding values for an identified dominant QTL with estimated polygenic breeding values using optimal versus standard allele substitution effects for the QTL, as a function of QTL gene frequency and for different levels of additive and dominance effects at the QTL. The QTL has two alleles B and b, with genotypic effects $a\sigma$, $d\sigma$ and $-a\sigma$ for BB, Bb and bb, where σ is the standard deviation of estimates of polygenic breeding values. Fractions selected are 0.20 for males and females.

Figure 2 compares the response to one generation of optimal GAS to response to standard GAS, as a function of frequency of the favorable allele at the QTL. The results are shown for varying levels of additive and dominance effects at the QTL. QTL effects are expressed relative to the standard deviation of EBV for

polygenic effects (σ), which is what determines the selection response for the QTL for polygenes, rather than relative to the genetic or phenotypic standard deviation. Therefore, the results in *figure 2* hold for specified magnitudes a and d in terms of standard deviations of EBV but regardless of heritability and phenotypic standard deviations. Relative QTL effects in *figure 2* can, however, be converted to values relative to the genetic standard deviation by multiplying a and d by the accuracy of EBV and to values relative to the phenotypic standard deviation by multiplying a and d by accuracy and the square root of heritability. For example, with polygenic EBV based on own phenotype alone for a trait with heritability equal to 0.25, and a phenotypic standard deviation equal to one, one standard deviation of EBV converts to 0.5 genetic standard deviations (accuracy = 0.5) and to 0.25 phenotypic standard deviations (square root of heritability = 0.5). Hence, a QTL with $a = 1\sigma$ represents a gene with only moderate effects for a trait with low heritability. For *figure 2*, the standard deviation and accuracy of EBV is assumed equal for males and females.

Over a single generation, benefits of optimal GAS over standard GAS were the greatest for QTL frequencies between 0.3 and 0.5 and increased with the magnitude of additive and dominance effects at the major gene (*figure 2*). Extra response was negligible for gene frequencies less than 0.05 and greater than 0.85. Extra response was greater than 8 % for QTL with large effects ($a > 1\sigma$) and complete dominance ($d > 1\sigma$) and with the favorable allele at intermediate frequency. For several combinations of parameter values, extra responses showed bi-modality as a function of gene frequencies.

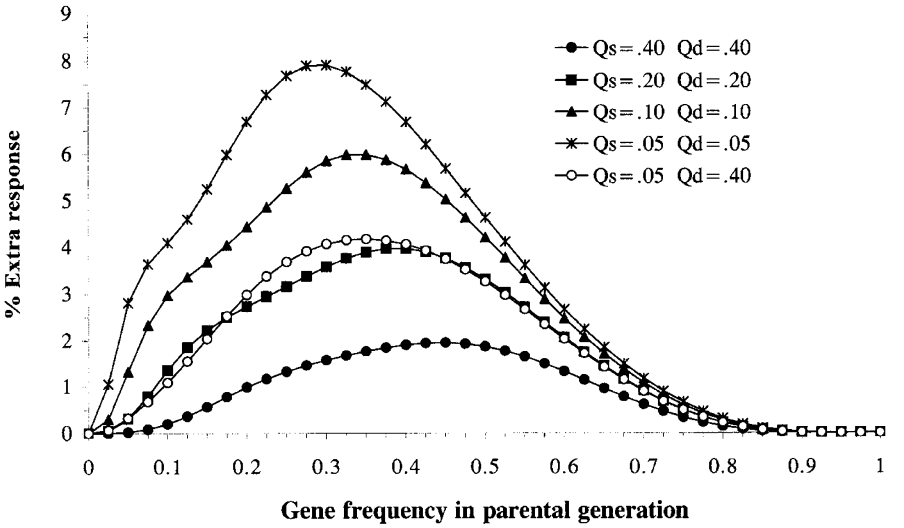

Figure 3. Extra (%) response from one generation selection on an index that combines breeding values for an identified dominant QTL with estimates of polygenic breeding values using optimal versus standard allele substitution effects for the QTL, as a function of QTL gene frequency and for different fractions selected among males (Qs) and females (Qd). The QTL has alleles B and b, with genotypic effects $+\sigma$, $+\sigma$ and $-\sigma$ for BB, Bb and bb, where σ is the standard deviation of estimates of polygenic breeding values.

Figure 3 shows the effect of selection intensity on extra response from optimal selection for a QTL with complete dominance and $a = 1\sigma$. The benefit of optimal selection increased with the intensity of selection. Selection of 5 % among males and 40 % among females had similar results as selection of 20 % for both males and females.

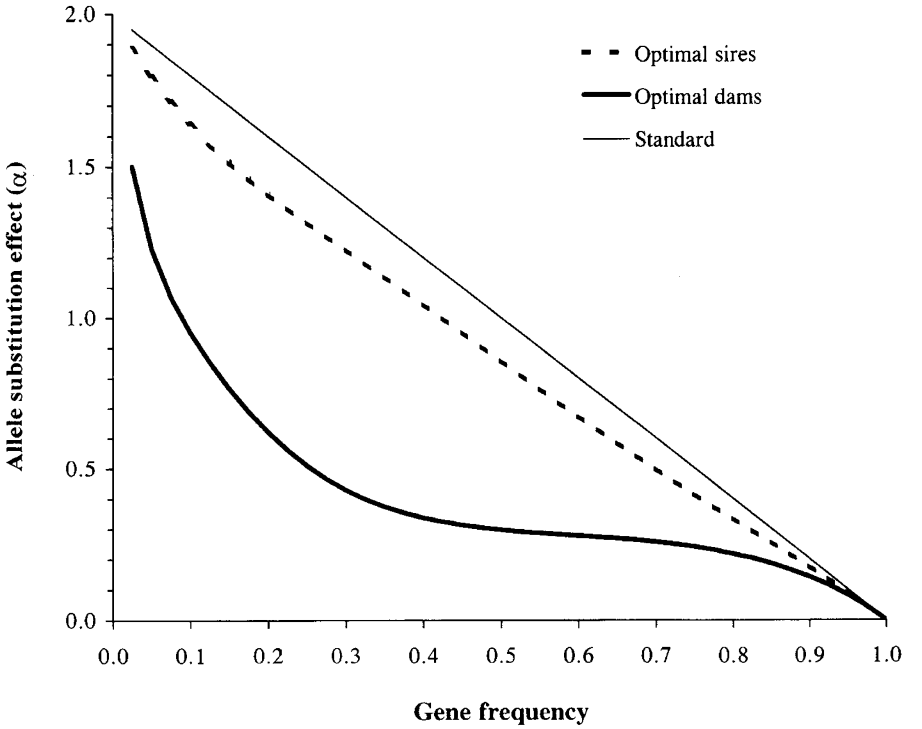

Figure 4. Standard and optimal allele substitution effects for an identified QTL with complete dominance as a function of QTL gene frequency, with 5 and 40 % of sires and dams selected on an index that combines QTL breeding values and polygenic breeding value estimates. The QTL has alleles B and b, with genotypic effects $+\sigma$, $+\sigma$ and $-\sigma$ for BB, Bb and bb, where σ is the standard deviation of estimates of polygenic breeding values.

Figure 4 shows the relationship between optimal allele substitution effects at the QTL and gene frequency for a QTL with complete dominance and with 5 % selected among males and 40 % among females. The standard allele substitution effect changes with gene frequency in a linear manner, based on $\alpha = a+(q-p)d$. Optimal allele substitution effects changed in a nonlinear manner, depending on QTL frequency among mates. Optimal allele substitution effects were lower than the standard substitution effects. For females, optimal substitution effects were up to 75 % lower than standard substitution effects. Optimal allele substitution effects were more greatly affected for females than males because selection intensity was greater for males, and, therefore, QTL frequency differed more drastically from QTL frequency among all candidates for selected males than for selected females. For recessive QTL (negative dominance), an opposite

effect would occur (results not shown); optimal breeding values are greater than standardized breeding values under selection because breeding values $(a + (q - p)d)$ increase with p for negative d. This increase in QTL breeding values will increase the emphasis on QTL relative to polygenes.

3.2. Multiple generation response

Responses to optimal and standard GAS were also compared over multiple generations, starting from a population in gametic phase equilibrium. QTL allele substitution effects were updated each generation for both optimal and standard GAS to account for the changes in gene frequency and gametic phase disequilibrium between the QTL and polygenes. Polygenic means by genotype class were assumed known without error. Polygenic variance was assumed to remain constant.

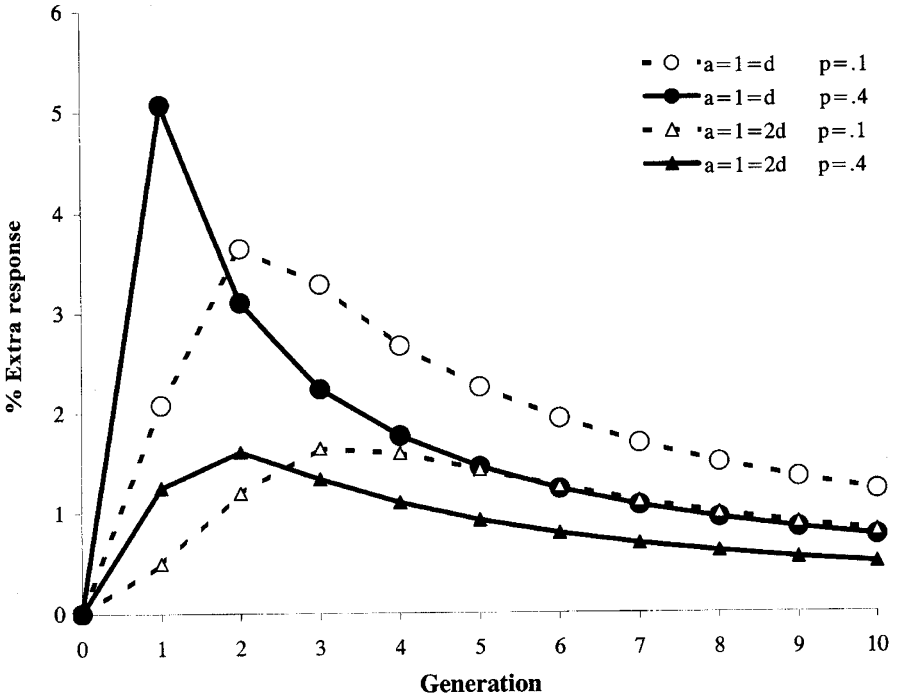

Figure 5. Extra (%) cumulative response from selection on an index that combines breeding values for an identified dominant QTL with polygenic breeding value estimates using one generation optimal versus standard allele substitution effects for the QTL, for different QTL effects and initial gene frequencies (p). The QTL has alleles B and b, with genotypic effects $a\sigma$, $d\sigma$ and $-a\sigma$ for BB, Bb and bb, where σ is the standard deviation of estimates of polygenic breeding values. Fractions selected are 0.20 for males and females.

Figure 5 shows the extra cumulative benefit of selection on optimal over standard QTL breeding values. *Figure 6* shows changes in gene frequency for the two selection strategies. Cumulative benefits increased over generations

until gene frequencies were between 0.3 and 0.5 and then decreased. This trend is consistent with the relationship between single generation response and gene frequency observed in *figures 2* and *3*. Extra cumulative responses after ten generations were relatively small (less than 2 % for the chosen examples).

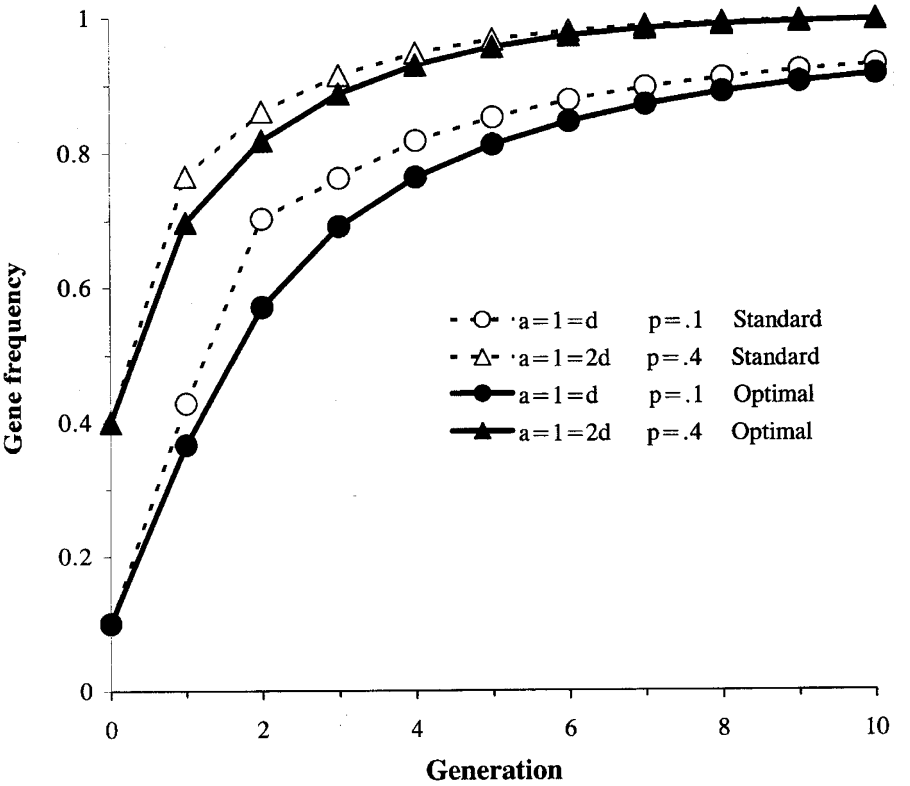

Figure 6. Change in gene frequency at a QTL from selection on an index that combines breeding values for an identified QTL with polygenic breeding value estimates using one generation optimal or standard allele substitution effects for the QTL, for different QTL effects and initial QTL gene frequencies (p). The QTL has alleles B and b, with genotypic effects $a\sigma$, $d\sigma$ and $-a\sigma$ for BB, Bb and bb, where σ is the standard deviation of polygenic breeding value estimates. Fractions selected are 0.20 for sires and dams.

4. DISCUSSION

The objective of this paper was to derive breeding values for a single locus that, when used in combination with EBV for polygenic effects, maximize single generation response to selection based on expected performance of progeny. Single locus breeding values thus derived were equivalent to breeding values derived on the standard quantitative genetic theory [4] but with the average effect of allele substitution, α derived from gene frequency among mates, rather than frequency in the unselected parental generation. With $\alpha = a + (q - p)d$,

the difference between optimal and standard breeding values for an individual locus therefore depends on the degree of dominance, d, and the effect of selection on the gene frequency among selected mates. The latter depends on selection emphasis that is placed on the individual locus and its effect and frequency. With phenotypic selection and when the trait is affected by a large number of genes of minor effect, the effect of selection on gene frequency will be small, and, hence, the difference between optimal and standard breeding values for a single locus will be minimal. With direct selection on QTL of sizeable effect, selection can, however, have a substantial impact on gene frequencies, and, therefore, optimal QTL breeding values can differ significantly from standard breeding values for a locus with dominance. This is illustrated in *figure 4*. The importance of derivation of optimum breeding values for a single locus lies in the current advances in molecular genetics, which lead to the uncovering of loci that affect quantitative traits, either by direct identification or indirectly through linked genetic markers. Use of this information in genetic improvement involves combining information on identified QTL with EBV for the collective effects of other genes that affect the trait (polygenic effects). The results from this paper show that, if the QTL exhibits dominance, substantial additional genetic progress can be made over a single generation if breeding values for the QTL take into account the effect of selection on gene frequencies among mates. Although benefits were small for QTL with moderate additive and dominance effects, improvements of up to 9 % in single generation response were observed for QTL with larger additive and dominance effects (see *figure 2*). Greatest benefits for the use of optimal over standard QTL breeding values were obtained for gene frequencies in the parental generation between 0.3 and 0.5, depending on the magnitude of the QTL effects. For a QTL with positive dominance, genetic variance contributed by the QTL and, therefore, the opportunity to change gene frequency is greatest for this range of gene frequencies [4].

The use of optimal QTL breeding values over successive generations resulted in greater cumulative response than the use of standard QTL breeding values (*figure 5*), although the benefit of optimal over standard breeding values decreased over generations. It is important to note that the optimal QTL breeding values derived here maximize single generation responses but may not maximize cumulative response over multiple generations. This has been illustrated by several authors (e.g. [7, 14]) for additive genes, for which standard QTL breeding values maximize single generation response, and by others (e.g. [11]) for QTL with dominance. The reason for the suboptimality of QTL selection strategies that maximize single generation response over multiple generation is that selection changes not only the population mean but also population parameters (frequency and, thereby, variance at the QTL) [2]. Single generation selection thereby affects opportunities for response in subsequent generations. Manfredi et al. [12] and Dekkers and Van Arendonk [2] developed methods to optimize QTL selection over multiple generations. The additional benefit of multiple generation optimization over single generation optimization will be investigated in subsequent work.

In the present study, QTL genotypes could be observed directly, and the effect of the QTL was assumed known without error. In many cases, QTL genotype must be inferred from linked genetic markers, and QTL effects will be estimated with some error. Both these factors will reduce the effect of selection

on changes in frequency at the QTL and, therefore, the difference between optimal and standard breeding values. With uncertainty about estimates of QTL effects, the effect of selection on QTL frequencies may be difficult to predict. This will increase the errors of prediction of optimal breeding values. It must also be noted that derivation of optimal QTL breeding values requires estimates of additive (a) and dominance (d) effects at the QTL, as well as an estimate of the frequency of the QTL. These estimates may be difficult to obtain in outbred populations based on linked markers. For example, the best linear unbiased prediction method developed by Fernando and Grossman [5] and extended by others for the incorporation of marker information in breeding value estimation estimates the average effect of the QTL, rather than separate additive and dominance effects. For non-additive QTL, the resulting QTL breeding value estimates will depend on the QTL frequency among mates of animals that contributed information to estimate the QTL effect. With selection on the QTL, the QTL frequency among mates of these animals may not be the same as the QTL frequency among individuals to which animals that are selected based on the QTL will be mated. The same holds for the multiple regression methods suggested by Lande and Thompson [10], in which marker effects are estimated as linear coefficients of regression of phenotypes on number of marker alleles. Implementation of optimal QTL breeding values in strategies for marker-assisted selection in outbred populations, therefore, requires further investigation.

ACKNOWLEDGMENTS

Financial support from the Iowa Pork Producers Association through the National Pork Board for aspects of this work is greatly appreciated. This is Journal Paper No. J-18323 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, USA (Project No. 3456) and supported by the Hatch Act and State of Iowa funds.

REFERENCES

[1] Bulmer M.G., The Mathematical Theory of Quantitative Genetics, Clarendon Press, Oxford, 1980.

[2] Dekkers J.C.M., Van Arendonk J.A.M., Optimizing selection for quantitative traits with information on an identified locus in outbred populations, Genet. Res. (1998) 257–275.

[3] Ducrocq V., Quaas R.L., Prediction of genetic response to truncation selection across generations, J. Dairy Sci. 71 (1988) 2543–2553.

[4] Falconer D.S., MacKay T.F.C., Introduction to Quantitative Genetics, Longman, Harlow, UK, 1996.

[5] Fernando R.L., Grossman M., Marker-assisted selection using best linear unbiased prediction, Genet. Sel. Evol. 21 (1989) 467.

[6] Gibson J.P., An Introduction to the Design and Economics of Animal Breeding Strategies, Course notes, Centre for Genetic Improvement of Livestock, University of Guelph, Canada, 1995, p. 149.

[7] Gibson J.P., Short-term gain at the expense of long-term response with selection of identified loci, Proc. 5th World Congr. Genet. Appl. Livest. Prod. 21 (1994) 201–204.

[8] Hazel L.N., The genetic basis of constructing selection indexes, Genetics 28 (1943) 476–490.

[9] Kennedy B.W., Quinton M., Van Arendonk J.A.M., Estimation of effects of single genes on quantitative traits, J. Anim. Sci. 70 (1992) 2000–2012.

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

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

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

[13] Neimann-Sorensen A., Robertson A., The association between blood groups and several production characteristics in three Danish cattle breeds, Acta Agric. Scand. (1961) 161–196.

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

[15] Smith C., Improvement of metric traits through specific genetic loci, Anim. Prod. 9 (1967) 349–358.

[16] Smith C., Webb A.J., Effects of major genes on animal breeding strategies, Z. Tierzucht. Zuchtgsbiol. 98 (1981) 161–169.

[17] Soller M., The use of loci associated with quantitative effects in dairy cattle improvement, Anim. Prod. 27 (1978) 133–139.

APPENDIX I: Bisection method to determine unique truncation point to select across multiple Normal distributions

Selection of a fraction Q by truncation across three distributions with frequencies p_i, mean μ_i $(i = 1, 2, 3)$ and standard deviation σ_i. Let c be the unique truncation point on the original scale and x_i and f_i the standardized truncation point and fraction selected for distribution i. Based on the definition of a standardized truncation point, $x_i = (c - \mu_i)/\sigma_i$ and $f_i = 1 - \Phi(x_i)$, where Φ is the cumulative Normal distribution function.

Then, truncation point c must be chosen such that $p_1 f_1 + p_2 f_2 + p_3 f_3 = Q$.

The following iterative procedure can be used to find truncation point c (based on [6]).

1) For all i, find the standardized truncation point x_i corresponding to $1 - \Phi(x_i) = Q$ using the inverse Normal distribution function.

2) Convert standardized truncation points x_i to the original scale based on $c_i = x_i\sigma_i + \mu_i$. Choose the lowest c_i as lower bound for c (c_L) and the highest c_i as the upper bound for c (c_U) (c must lie between c_L and c_U).

3) Compute the midpoint between c_L and c_U: $c_M = (c_L + c_U)$.

4) Compute standardized truncation points corresponding to c_M for each distribution: $x_i = (c_M - \mu_i)/\sigma_i$ and the corresponding proportions selected: $f_i = 1 - \Phi(x_i)$.

5) Compute the total proportion selected as: $Q_M = p_1 f_1 + p_2 f_2 + p_3 f_3$.

6) If $|Q_M - Q| <$ convergence criterion, the unique truncation point has been found: $c = c_M$

7) If $Q_M - Q < 0$, then set $c_U = c_M$. If $Q_M - Q > 0$, then set $c_L = c_M$. Return to step 3.