

# Optimizing selection for quantitative traits with information on an identified locus in outbred populations

J. C. M. DEKKERS<sup>1\*</sup> AND J. A. M. VAN ARENDONK<sup>2</sup>

<sup>1</sup>Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, ON N1G 2W1, Canada

<sup>2</sup>Animal Breeding and Genetics Unit, Wageningen Institute of Animal Sciences, PO Box 338, 6700 AH Wageningen, the Netherlands

(Received 15 July 1997 and in revised form 20 November 1997)

## Summary

Methods to formulate and maximize response to selection for a quantitative trait over multiple generations when information on a quantitative trait locus (major gene) is available were developed to investigate and optimize response to selection in mixed inheritance models. Deterministic models with and without gametic phase disequilibrium between the major gene and other genes that affect the trait (polygenes) were considered. Genetic variance due to polygenes was assumed constant. Optimal control theory was used to formulate selection on an index of major gene effects and estimates of polygenic breeding values and to derive index weights that maximize cumulative response over multiple generations. Optimum selection strategies were illustrated using an example and compared with mass selection and with selection with full emphasis on the major gene (genotypic selection). The latter maximizes the single-generation response for a major gene with additive effects. For the example considered, differences between selection methods in cumulative response at the end of a planning horizon of 5, 10, or 15 generations were small but responses were greatest for optimum selection. Genotypic selection had the greatest response in the short term but the lowest response in the longer term. For optimum selection, emphasis on the major gene changed over generations. However, when accounting for variance contributed by the major gene, optimum selection resulted in approximately constant selection pressure on the major gene and polygenes over generations. Suboptimality of genotypic selection in the longer term was caused not so much by gametic phase disequilibrium but rather by unequal selection pressure on the major gene (and, therefore, on polygenes) over generations, as frequency and variance at the major gene changed. Extension of methods to more complex breeding structures, genetic models and objective functions is discussed.

## 1. Introduction

Current breeding programmes for quantitative traits in livestock involve selection of parents on estimates of their breeding values, without knowledge of the individual's genotype for individual genes. Typically, breeding values are estimated on the basis of phenotypic information on the animal itself and its relatives, using Best Linear Unbiased Prediction (BLUP) procedures (Henderson, 1988). Recent developments in molecular genetics are, however, leading to the uncovering of individual genes that have an effect on quantitative traits (quantitative trait loci, or QTLs)

and of genes (genetic markers) that are closely linked to QTLs. Use of information on identified QTLs in breeding programmes, along with the traditional phenotypic information, can lead to enhanced rates of genetic improvement, especially in cases where phenotypic information is not available on selection candidates or expensive to collect, or if the trait has low heritability (Smith & Simpson, 1986).

Optimum use of information on identified QTLs in selection programmes requires development of selection criteria that combine information from single genes with phenotypic information. Principles behind statistical procedures to compute such selection criteria have been developed based on BLUP (reviewed in Van Arendonk *et al.*, 1994). However, Gibson

\* Corresponding author. Department of Animal Science, 201 Kildee Hall, Iowa State University, Ames, IA 50011, USA.

(1994) showed that, although such selection criteria can maximize genetic progress in the short term (i.e. in the current generation), they may not maximize response to selection in the longer term. In fact, Gibson (1994) found that traditional selection, based on phenotypic information alone, resulted in greater genetic improvement in the longer term than selection on a combination of phenotypic information and information on identified genes (genotypic selection). Thus, selection criteria that are optimal in the short term may not lead to maximum response in the longer term. Similar results were found by Woolliams & Pong-Wong (1995) with selection on a major gene and by Ruane & Colleau (1995) with BLUP selection on genetic markers linked to a major gene. For sex-limited traits, Van der Beek & Van Arendonk (1994) and Ruane & Colleau (1996) found use of information from genetic markers linked to a major gene to result in greater response than selection without information from genetic markers, regardless of the length of the planning horizon. This does not, however, mean that the BLUP selection criterion used in these studies maximized responses to selection.

Loss of longer-term response with genotypic selection or with BLUP selection on genetic markers linked to a QTL is caused by a reduction in polygenic response, which can be attributed to a reduction in effective selection intensity that is applied to polygenes when information from the major gene is considered (Gibson, 1994). The reason why, in the longer term, loss in polygenic response is not offset by increased response for the major gene is unclear; Woolliams & Pong-Wong (1995) suggested the build-up of gametic phase disequilibrium between the major gene and polygenes (Kennedy *et al.*, 1992) as possible cause. However, Ruane & Colleau (1995) found little difference in gametic phase disequilibrium for a marked QTL and polygenes when comparing marker-assisted selection with selection based on BLUP of breeding values without marker information. Ruane & Colleau (1995) suggested reduced accuracy of estimation of polygenic breeding values with use of marker information as the reason for the reduced polygenic response with marker-assisted selection. This does not, however, explain the results for selection on a major gene with known effect. In the present study, the relationship between frequency of the major gene and genetic variance contributed by the major gene is suggested and investigated as another possible reason for loss in long-term response to genotypic selection.

Objectives of this study were to develop a theoretical framework for methods to formulate and optimize selection for quantitative traits with information on an identified QTL, to derive selection criteria that maximize cumulative response within a given planning horizon, and to investigate the reason and nature of

losses in longer-term response with genotypic selection. In this study, methods are developed and illustrated for a simple breeding structure, selection strategy and genetic model, with discrete generations and equal selection in both sexes, to allow illustration of concepts. A major gene with additive effects is considered. Extensions to more complex (and realistic) situations are discussed.

Methodology to optimize selection with information from major genes developed herein is based on optimal control theory (Bryson & Ho, 1975; Kamien & Schwartz, 1981; Lewis, 1986). Optimal control theory is used extensively in economics and engineering to formulate and optimize decision problems that involve multiple stages and in which transition of the system from stage  $t$  to stage  $t+1$  can be described by first-order difference equations. The latter implies that transition of the system from stage  $t$  to  $t+1$  depends only on the state of the system in stage  $t$  and on the decisions made at stage  $t$  and not on *how* the system reached the state in stage  $t$ . Dekkers *et al.* (1995) used optimal control theory to formulate and solve optimization of selection for non-linear profit functions over multiple generations and suggested use of optimal control as a tool to solve other multiple-generation selection problems in animal breeding.

## 2. Methods

### (i) Assumptions and general principles

Consider a population of infinite size with discrete generations, selection of a fraction  $Q$  of males and females to be parents of the next generation, and random mating of selected parents. Selection is for a quantitative trait that is affected by a major gene with additive effects and by additive polygenic effects. The major gene has two alleles (B and b) with additive effects and frequencies  $p_t$  and  $1-p_t$  in generation  $t$ . Genotypes BB, Bb and bb are denoted by  $m = 1, 2$  and 3, and have genotypic values denoted by  $g_m$ , where  $g_1 = a$ ,  $g_2 = 0$  and  $g_3 = -a$ . Genotypes at the major gene are available on all individuals prior to the age of selection. Polygenic breeding values follow a Normal distribution. Heritability for polygenic breeding values is equal to  $h_t^2$  in generation  $t$ . The average breeding value of the population in generation  $t$  is equal to  $\bar{G}_t = a(2p_t - 1) + \bar{A}_t$ , where  $\bar{A}_t$  is the average polygenic value of animals in generation  $t$ . This relationship holds with or without gametic phase disequilibrium between the major gene and polygenes.

In general, selection in generation  $t$  is by truncation selection on an index that combines the value of the major genotype ( $g_m$ ) with an estimate of the polygenic breeding value for animal  $i$  of major genotype  $m$  ( $\hat{A}_{imt}$ ):

$$I_{imt} = b_{mt}(g_m + \bar{A}_{mt} - \bar{A}_{2t}) + \hat{A}_{imt} \quad (1)$$

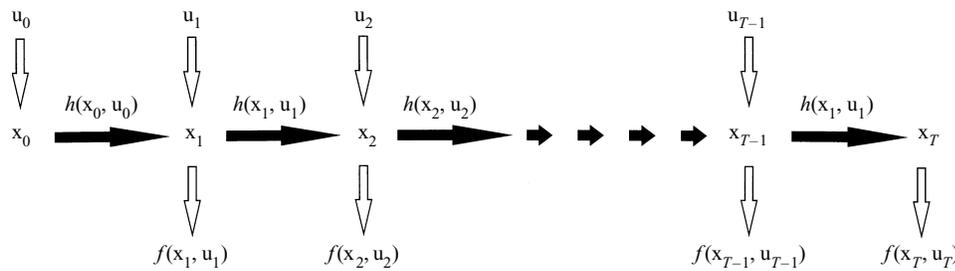


Fig. 1. Schematic presentation of a multiple-stage decision process over  $T$  stages with control variables ( $u_t$ ), state variables ( $x_t$ ), state equations ( $h(x_t, u_t)$ ) and stage output ( $f(x_t, u_t)$ ). Maximization of a function of stage outputs over a planning horizon involves finding optimum control variables at each stage subject to the state equations at each stage and given the value of state variables at stage 0.

where  $I_{imt}$  is the selection criterion for animal  $i$  of major genotype  $m$  in generation  $t$ ,  $\hat{A}_{imt}$  is an estimate of the polygenic breeding value for animal  $i$  with genotype  $m$  in generation  $t$ , as a deviation from the average value of animals with genotype  $m$  in generation  $t$  ( $= g_m + \bar{A}_{mt} - \bar{A}_{2t}$ ),  $b_{mt}$  is the weight put on the average value of genotype  $m$  in generation  $t$  and  $\bar{A}_{mt}$  is the average polygenic value of animals with major genotype  $m$  in generation  $t$ . In a large population, the average value of each major genotype in generation  $t$  ( $g_m + \bar{A}_{mt} - \bar{A}_{2t}$ ) that is required in index (1) can be estimated with small error based on contrasts between genotypes. Estimation without error is assumed here. The term  $\bar{A}_{mt} - \bar{A}_{2t}$  in (1) represents the extent of gametic phase disequilibrium between the major gene and polygenes. Without gametic phase disequilibrium, (1) simplifies to:  $I_{imt} = b_{mt}g_m + \hat{A}_{imt}$ .

Note that for  $m = 2$ , (1) simplifies to  $I_{2t} = \hat{A}_{imt}$ . Weights  $b_{2t}$  are, therefore, immaterial. Also note that in this formulation, estimates of polygenic breeding values  $\hat{A}_{imt}$  can be based on BLUP, incorporating information from relatives, with a model that includes major genotype as a fixed effect (e.g. Kennedy *et al.*, 1992). With mass selection,  $b_{mt} = h_t^2$  for all  $t$ , and  $\hat{A}_{imt} = h_t^2(P_{imt} - g_m - \bar{A}_{mt} + \bar{A}_{2t})$ , where  $P_{imt}$  is phenotype. With genotypic selection,  $b_{mt} = 1$  for all  $t$ .

Based on the above formulation, maximization of cumulative genetic response to selection in generation  $t = T$  then involves solution of a multiple-stage decision problem (Lewis, 1986) in which index weights  $b_{mt}$  must be optimized for each generation  $t$  ( $t = 0$  to  $T - 1$ ), in order to maximize the objective function:

$$\bar{G}_T = a(2p_T - 1) + \bar{A}_T. \tag{2}$$

This multiple-stage decision problem is illustrated in Fig. 1 and can be formulated and solved using optimal control theory (Lewis, 1986). This involves definition of state variables, control variables, state equations and an objective function. State variables describe the state of the system at each stage (generation)  $t$  and in our case include frequency of the major gene, average polygenic breeding values, and genetic variances. Control variables are the decision

variables that are under the control of the decision maker. In our case, control variables are the index weights  $b_{mt}$  for each generation  $t$ . State equations, which are formulated for each generation  $t$ , describe the transition of the state variables from generation  $t$  to generation  $t + 1$ . For the purposes of optimal control, state equations must be functions of state and control variables in the previous generation only (first-order difference equations). The objective function must be a separable function of output of each stage (Lewis, 1986). In our case, only output from the final generation is considered in the objective function but more complex objective functions (e.g. cumulative discounted response) can be considered also.

In principle, truncation selection on index  $I_{imt}$  (equation (1)) in a given generation  $t$  involves truncation selection across three Normal distributions with means  $b_{mt}(g_m + \bar{A}_{mt} - \bar{A}_{2t})$ , and standard deviation  $\sigma_{mt}$ , where  $\sigma_{mt}$  is the standard deviation of estimates of polygenic breeding values  $\hat{A}_{imt}$ . As indicated previously, estimates of polygenic breeding values can be based on BLUP, incorporating information from relatives. The truncation selection process across the three major genotypes is illustrated in Fig. 2 for  $\sigma_{mt} = \sigma$ . In Fig. 2,  $x_{mt}$  is the standard Normal truncation point for genotype  $m$  and  $f_{mt}$  is the proportion selected from genotype  $m$ . When expressed on the scale of estimated breeding values and as a deviation from the mean, truncation points are equal to  $\sigma_{mt}x_{mt}$ . With truncation selection across genotypes, differences between truncation points for the three genotypes are equal to differences between means of the distributions. Means are equal to  $b_{mt}(g_m + \bar{A}_{mt} - \bar{A}_{2t})$  when selecting on index (1). Rearranging gives the following relationships between weights  $b_{1t}$  and  $b_{3t}$  of index (1) and truncation points  $x_{mt}$ :

$$b_{1t} = (\sigma_{2t}x_{2t} - \sigma_{1t}x_{1t}) / (a + \bar{A}_{1t} - \bar{A}_{2t}), \tag{3a}$$

$$b_{3t} = (\sigma_{3t}x_{3t} - \sigma_{2t}x_{2t}) / (a + \bar{A}_{2t} - \bar{A}_{3t}). \tag{3b}$$

With  $\sigma_{mt} = \sigma$ , (3a) and (3b) simplify to

$$b_{1t} = \sigma(x_{2t} - x_{1t}) / (a + \bar{A}_{1t} - \bar{A}_{2t}), \tag{4a}$$

$$b_{3t} = \sigma(x_{3t} - x_{2t}) / (a + \bar{A}_{2t} - \bar{A}_{3t}). \tag{4b}$$

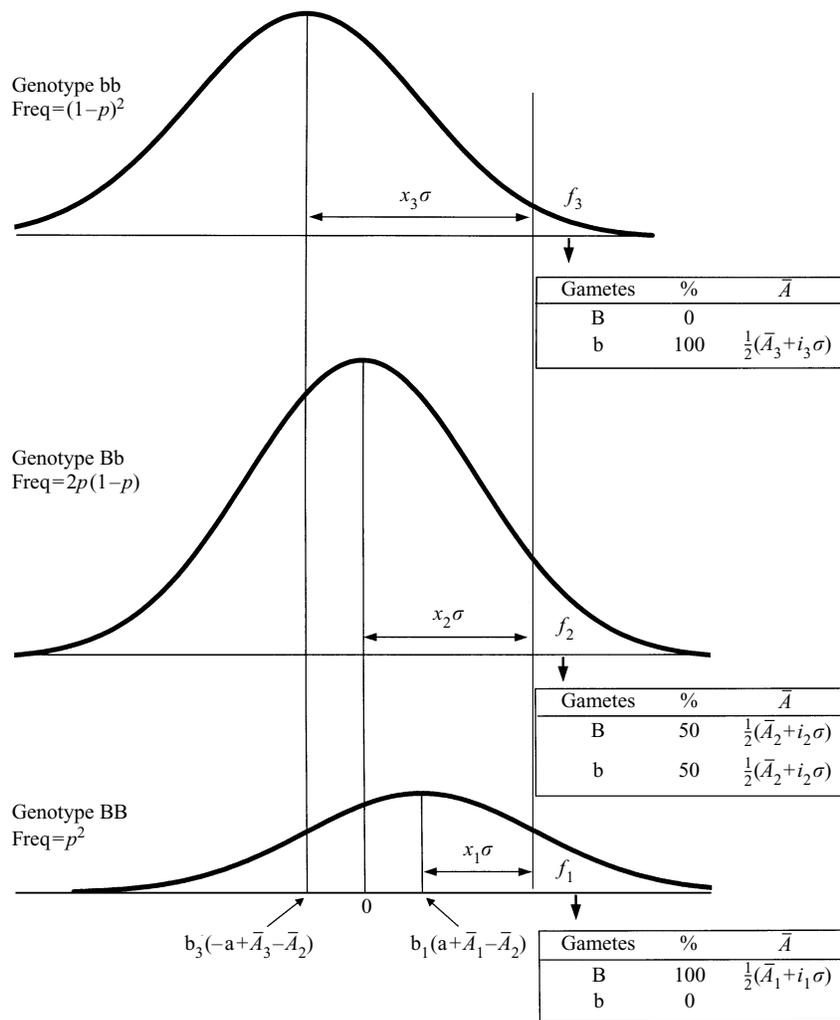


Fig. 2. Schematic presentation of selection on an index of major genotype value and estimates of polygenic breeding values (EBV) in the form of truncation selection across Normal distributions of estimates of polygenic breeding values corresponding to the three major genotypes (BB, Bb and bb). The major gene has frequency  $p$  and additive genetic value  $a$ . Parameter  $\sigma$  is the standard deviation of estimates of polygenic breeding values.  $\bar{A}_i$  is the average polygenic value of animals with major genotype  $i$  ( $i = 1, 2, 3$  for major genotypes BB, Bb and bb, respectively);  $x_i$ ,  $f_i$  and  $i_i$  are, respectively, the standardized truncation point, fraction selected and selection intensity for animals with major genotype  $i$ . Parameters  $b_1$  and  $b_3$  are weights on major genotype value in the index for animals with major genotype BB and bb, respectively.

Based on the above, index weights  $b_{mt}$  are uniquely related to truncation points  $x_{mt}$  and, thereby, to proportions selected ( $f_{mt}$ ). Optimization of weights  $b_{mt}$  is therefore equivalent to optimizing truncation points  $x_{mt}$  and equivalent to optimizing proportions selected from the three distributions of EBV (Fig. 2) for each generation  $t$ . Therefore, variables  $f_{mt}$  can be used as control variables in the optimal control formulation instead of  $b_{mt}$ , which is the approach used in what follows. Optimum solutions for  $f_{mt}$  are then converted to optimum solutions for  $b_{mt}$  based on (4).

In this paper, gametic phase disequilibrium for polygenic effects will be assumed absent and, therefore,  $\sigma_{mt} = \sigma$  for all  $m$  and  $t$ . Initially, deterministic models in which the major gene and polygenes are in gametic phase equilibrium in the progeny generation will be

investigated. This model is considered here to elucidate the factors that affect suboptimality of genotypic selection and characteristics of optimal selection strategies. The assumption of gametic phase disequilibrium will subsequently be relaxed.

(ii) *No gametic phase disequilibrium between major gene and polygenes*

In the model without gametic phase disequilibrium between the major gene and polygenes, gametic phase disequilibrium generated in parents through the process of selection was assumed to be completely resolved during meiosis. The average polygenic values of B and b gametes were, therefore, equal to half the pooled average polygenic value of selected parents

and average polygenic values were equal for all three progeny genotypes ( $\bar{A}_{1t} = \bar{A}_{2t} = \bar{A}_{3t} = \bar{A}_t$ ).

Based on the above and the principles of response to selection, the problem of maximizing  $\bar{G}_T$  (equation (2)) can then be formulated as an optimum control problem (Dekkers *et al.*, 1995; Lewis, 1986), with  $p_t$  and  $\bar{A}_t$  as state variables, and  $f_{mt}$  ( $m = 1, 2, 3$ ) as control variables:

$$\text{Max}_{f_{mt}} \{a(2p_T - 1) + \bar{A}_T\} \tag{5}$$

subject to:

$$f_{1t}p_t^2 + f_{2t}2p_t(1-p_t) + f_{3t}(1-p_t)^2 = Q \tag{5a}$$

for  $t = 0$  to  $T-1$ ,

$$p_{t+1} = \frac{1}{Q} \{f_{1t}p_t^2 + f_{2t}2p_t(1-p_t)\} \tag{5b}$$

for  $t = 0$  to  $T-1$ ,

$$\begin{aligned} \bar{A}_{t+1} &= \bar{A}_t + \frac{\sigma}{Q} \{p_t^2 f_{1t} i_{1t} + 2p_t(1-p_t) f_{2t} i_{2t} \\ &\quad + (1-p_t)^2 f_{3t} i_{3t}\} \\ &= \bar{A}_t + \frac{\sigma}{Q} \{p_t^2 z_{1t} + 2p_t(1-p_t) z_{2t} + (1-p_t)^2 z_{3t}\} \end{aligned} \tag{5c}$$

for  $t = 0$  to  $T-1$ ,

and given  $\bar{A}_0, p_0$  and  $Q$ .

In this formulation, (5a) are the constraints on the overall fraction selected ( $Q$ ) within each generation. Frequencies of major genotypes in a given generation are those following Hardy–Weinberg equilibrium (Falconer & Mackay, 1996), which holds with equal selection in males and females and random mating of selected parents. Equations (5b) and (5c) are state equations for frequency of the major gene and average polygenic breeding values, respectively. Equation (5c) represent the single-generation response to selection in polygenic breeding values. These are based on pooled selection differentials within each major genotype class. In (5c),  $z_{mt}$  is the height of the standard Normal distribution at the standardized truncation point  $x_{mt}$  for genotype class  $m$ , which results in a fraction  $f_{mt}$  of animals with genotype  $m$  to be selected. Note that  $z_{mt} = i_{mt} f_{mt}$  (Falconer & Mackay, 1996), where  $i_{mt}$  is the selection intensity corresponding to  $f_{mt}$ .

Using optimum control theory (Lewis, 1986), this multiple-stage decision problem can be solved by incorporating the constraint equations (5a), (5b) and (5c) into the objective function (5) using sets of Lagrange multipliers for each constraint equation:  $\epsilon_t$ ,  $\lambda_{t+1}$  and  $\gamma_{t+1}$ , for constraint equations (5a), (5b) and (5c), respectively, for  $t = 0$  to  $T-1$ . Lagrange multipliers are shadow values for the constraint equations, which can be interpreted as the marginal change in the objective function when the constraint

is relaxed by a marginal amount. Here, Lagrange multiplier  $\epsilon_t$  refers to the shadow value of the constraint on the fraction  $Q$  selected in generation  $t$ . Coefficients  $\lambda_{t+1}$  and  $\gamma_{t+1}$  refer to shadow values for the gene frequency and average polygenic breeding value attained in generation  $t+1$ .

After rearranging terms, incorporating the constraint equations into the objective function results in the following non-linear maximization problem:

$$\text{Max}_{f_{mt}} \{L\} \text{ given } \bar{A}_0, p_0 \text{ and } Q \tag{6}$$

with

$$L = \sum_{t=0}^{T-1} \{H_t - \lambda_t p_t - \gamma_t \bar{A}_t\} - \lambda_T p_T + \lambda_0 p_0 - \gamma_T \bar{A}_T + \gamma_0 \bar{A}_0 + a(2p_T - 1) + \bar{A}_T \tag{7}$$

with

$$\begin{aligned} H_t &= \frac{\lambda_{t+1}}{Q} \{f_{1t}p_t^2 + f_{2t}2p_t(1-p_t)\} \\ &\quad + \gamma_{t+1} \bar{A}_t + \frac{\sigma}{Q} [p_t^2 z_{1t} + 2p_t(1-p_t) z_{2t} + (1-p_t)^2 z_{3t}] \\ &\quad + \epsilon_t \{Q - f_{1t}p_t^2 - f_{2t}2p_t(1-p_t) - f_{3t}(1-p_t)^2\}. \end{aligned} \tag{8}$$

Within the context of optimal control theory (Lewis, 1986),  $H_t$  is referred to as the Hamiltonian function. Part of the rearrangement of terms that leads to (7) is such that  $H_t$  can be written exclusively as a function of variables that correspond to generation  $t$  and of variables that correspond to the constraints for generation  $t$  (i.e.  $\lambda_{t+1}$  and  $\gamma_{t+1}$ ). This formulation facilitates subsequent solution of the non-linear optimization problem based on its recursive properties.

Optimal solutions to (6) are derived by equating the first partial derivative of  $L$  with regard to each control variable, state variable and Lagrange multiplier to zero for each generation  $t$ . Resulting equations are given in the Appendix. Note from the Appendix that partial derivatives of  $L$  can be reduced to partial derivatives of  $H_t$ , which illustrates the utility of defining the Hamiltonian function. Manipulation of the resulting sets of equations, which is shown in the Appendix, results in two sets of recursive equations that must be met to attain the optimal solutions. The first set is a forward recursive set of equations in gene frequency  $p_t$  (equations (A 4f), which is identical to the set of constraint equations (5b)). This set of equations allows computation of  $p_{t+1}$  from  $p_t$  given  $f_{it}$  and has gene frequency in the initial generation ( $p_0$ ) as known starting value. The second set is a backward recursive set of equations for the standardized Normal truncation points (equations (A 10)). This set relates the difference in truncation points between major

genotypes in generation  $t-1$   $[(x_{3t-1}-x_{2t-1})$  and  $(x_{2t-1}-x_{1t-1})]$  to truncation points in generation  $t$  (as well as to the corresponding fractions selected and selection intensities). This set of recursive equations has  $(x_{3T-1}-x_{2T-1}) = (x_{2T-1}-x_{1T-1}) = a/\sigma$  as known starting value for generation  $T$  (see Appendix). Given the difference in truncation points, frequency  $p_t$ , and overall fraction selected  $Q$  (equation (5a)), truncation points  $x_{it}$  can be derived for each generation (see Appendix).

Although optimum solutions can not be derived analytically, these two sets of recursive equations, along with a constraint on the total fraction selected, as given by (5a), can be used to derive a numerical procedure to obtain the solution in an iterative manner. Such an iterative procedure, which is based on repeatedly using the forward recursive followed by the backward recursive equation, each time updating all variables involved, is given in the Appendix.

(iii) *With gametic phase disequilibrium between major gene and polygenes*

In the previous section, the major gene was assumed to be in gametic phase equilibrium with polygenes in each generation. Selection on a combination of major genotype value and polygenic breeding value, however, results in a negative association between major genotype and polygenic breeding values (Kennedy *et al.*, 1992). This negative association is due to the fact that parents selected from major genotype BB are selected with lower selection intensity for polygenic effects and have a lower average polygenic breeding value than parents selected from major genotypes Bb or bb, as illustrated in Fig. 2. This negative association or gametic phase disequilibrium can be modelled at the gametic level, as described below.

Let  $\bar{A}_{B,t}$  and  $\bar{A}_{b,t}$  be the average polygenic value of gametes that combine to produce animals for generation  $t$  and that contain major gene alleles B and b, respectively. The average polygenic value of animals in generation  $t$  with major genotype BB, Bb and bb is  $2\bar{A}_{B,t}$ ,  $\bar{A}_{B,t} + \bar{A}_{b,t}$  and  $2\bar{A}_{b,t}$ , respectively. Then, the overall average polygenic value in generation  $t$  is equal to

$$\bar{A}_t = 2p_t \bar{A}_{B,t} + 2(1-p_t) \bar{A}_{b,t} \tag{9}$$

With selection among animals in generation  $t$ , parents selected from major genotype class BB have average polygenic breeding value equal to  $2\bar{A}_{B,t} + i_{1t} \sigma$  and produce 100% B gametes with an average polygenic value equal to  $\bar{A}_{B,t} + \frac{1}{2}i_{1t} \sigma$  (Fig. 2). Similarly, parents with major genotype bb produce 100% b gametes with average polygenic value equal to  $\bar{A}_{b,t} + \frac{1}{2}i_{3t} \sigma$ . Parents with major genotype Bb produce 50% B gametes and 50% b gametes. When the major gene

and polygenes are unlinked, the average polygenic value of both types of gametes produced by Bb parents is equal to  $\frac{1}{2}(\bar{A}_{B,t} + \bar{A}_{b,t} + i_{2t} \sigma)$ .

The following recursive equation can then be set up for  $\bar{A}_{B,t}$ :

$$\bar{A}_{B,t+1} = \frac{f_{1t} p_t^2 (\bar{A}_{B,t} + \frac{1}{2} i_{1t} \sigma) + f_{2t} p_t (1-p_t) \frac{1}{2} (\bar{A}_{B,t} + \bar{A}_{b,t} + i_{2t} \sigma)}{f_{1t} p_t^2 + f_{2t} p_t (1-p_t)} \tag{10}$$

Note that (10) does not account for the fact that polygenic values of B gametes that produced generation  $t$  originated from two distinct distributions (BB and Bb parents) in generation  $t-1$  and, therefore, have a bi-modal distribution. In principle, however, these effects can be included in the model by defining extra genotype classes and corresponding state variables.

Realizing that the denominator of (10) is equal to  $p_{t+1} Q$  (equation (5b)), it is advantageous to introduce a new variable,  $W_{B,t} = p_t \bar{A}_{B,t}$ , for which the recursive equation simplifies to

$$W_{B,t+1} = \frac{1}{2Q} \{ [2f_{1t} p_t + f_{2t} (1-p_t)] W_{B,t} + f_{2t} p_t W_{b,t} + p_t \sigma [p_t z_{1t} + (1-p_t) z_{2t}] \} \tag{11a}$$

Similarly, the recursive equation for  $W_{b,t} = (1-p_t) \bar{A}_{b,t}$  is

$$W_{b,t+1} = \frac{1}{2Q} \{ [2f_{3t} (1-p_t) + f_{2t} p_t] W_{b,t} + f_{2t} p_t W_{B,t} + (1-p_t) \sigma [(1-p_t) z_{3t} + p_t z_{2t}] \} \tag{11b}$$

With  $p_t$ ,  $W_{B,t}$  and  $W_{b,t}$  as state variables, and  $f_{mt}$  ( $m = 1, 2, 3$ ) as control variables, the problem of maximizing cumulative response in generation  $T$  can then be formulated as an optimal control problem as follows:

$$\text{Max}_{f_{mt}} \{ a(2p_T - 1) + 2W_{B,T} + 2W_{b,T} \} \tag{12}$$

subject to (for  $t = 0$  to  $T-1$ )

$$f_{1t} p_t^2 + f_{2t} 2p_t (1-p_t) + f_{3t} (1-p_t)^2 = Q, \tag{12a}$$

$$p_{t+1} = \frac{1}{Q} \{ f_{1t} p_t^2 + f_{2t} p_t (1-p_t) \}, \tag{12b}$$

$$W_{B,t+1} = \frac{1}{2Q} \{ [2f_{1t} p_t + f_{2t} (1-p_t)] W_{B,t} + f_{2t} p_t W_{b,t} + p_t \sigma [p_t z_{1t} + (1-p_t) z_{2t}] \}, \tag{12c}$$

$$W_{b,t+1} = \frac{1}{2Q} \{ [2f_{3t} (1-p_t) + f_{2t} p_t] W_{b,t} + f_{2t} (1-p_t) W_{B,t} + (1-p_t) \sigma [(1-p_t) z_{3t} + p_t z_{2t}] \}, \tag{12d}$$

and given  $W_{B,0}$ ,  $W_{b,0}$ ,  $p_0$  and  $Q$ .

Similar to the situation without gametic phase disequilibrium, optimum solutions to the above non-linear optimization problem can be derived using principles of optimal control theory by incorporating the constraint equations into the objective function using Lagrange multipliers and setting equal to zero the first partial derivatives of the resulting function with respect to control variables, state variables and Lagrange multipliers. Manipulation of the resulting set of recursive equations results again in sets of recursive equations that can be used to formulate an iterative procedure to find the optimum truncation points  $x_{it}$  for each generation  $t$ . Derivations and an iterative procedure are given in the Appendix. Similar to the model without gametic phase disequilibrium, the solution involves iteration over a set of forward and a set of backward recursive equations. The forward recursive equations are in the state variables  $p_t$ ,  $W_{bt}$  and  $W_{bt}$  (equations (A 12i), (A 12g) and (A 12h) in the Appendix, which are equivalent to the constraint equations (12b), (12c) and (12d)). The backward recursive equations are in the Lagrange multipliers that correspond to each state variable (equations (A 14a), (A 14b) and (A 14c) in the Appendix). Within each iteration, truncation points  $x_{it}$  can be derived for each generation  $t$ , based on updated values for the state variables and Lagrange multipliers and given the constraint on the overall fraction selected (constraint equation (12a)). Iteration on these sets of recursive equations leads to the optimal solutions (see Appendix).

#### (iv) Example

To illustrate methods and allow an initial comparison of responses from mass selection, genotypic selection and optimum selection, albeit for a simple example, procedures were applied to selection of 20% per generation ( $Q = 0.2$ ) for a trait with an identified major gene with additive effect  $a = 0.25$  (no dominance) and frequency  $p_0 = 0.05$  in generation 0, a standard deviation of polygenic estimated breeding values of  $\sigma = 0.3$ , and average polygenic breeding values in generation 0 equal to  $\bar{A}_0 = \bar{A}_{B,0} = \bar{A}_{b,0} = 0$ . With regard to comparisons involving mass selection,  $\sigma = 0.3$  corresponds to the standard deviation of polygenic estimate of breeding values (EBV) based on one own record for a trait with  $\sigma_p = 1$  and  $h^2 = 0.3$  (the standard deviation of EBV for mass selection is equal to the standard deviation of  $h^2P$ , where  $P$  is phenotype, which is equal to  $h^2\sigma_p$ ). For genotypic and optimum selection, estimates of polygenic EBV are not restricted to use of own records only but polygenic EBV can represent BLUP EBV, incorporating information from relatives. Therefore, when comparing genotypic with optimum selection,  $\sigma$  refers to the standard

deviation of polygenic EBV that results from the process used for estimating polygenic breeding values, which would be the same for genotypic and optimum selection. Because the models for genotypic and optimum selection depend only on  $\sigma$ , there is no need to specify  $h^2$  explicitly when comparing genotypic and optimum selection and results apply to polygenic EBV that are estimated based on own phenotype, selection index or BLUP. When comparing genotypic or optimum selection with mass selection, however, polygenic EBV are assumed to be based on own phenotype only.

### 3. Results

#### (i) Mass selection versus genotypic selection

Broken lines in Fig. 3 show cumulative total response (major gene plus polygenes) to mass selection over 1 to 15 generations as a percentage of response to genotypic selection. Results are given for the model with (thick lines) and without (thin lines) consideration of gametic phase disequilibrium between the major gene and polygenes. Results confirm those of Gibson (1994) that genotypic selection gives greater cumulative response in initial generations but lower cumulative response in the longer term. However, results in Fig. 3 also indicate that the negative impact of genotypic selection on the longer-term response is present also for the model in which gametic phase disequilibrium between the major gene and polygenes was not included. Therefore, reduced longer-term response to genotypic selection is not caused solely by a build-up of gametic phase disequilibrium between the major gene and polygenes.

Comparing relative responses to mass and genotypic selection with and without gametic phase disequilibrium (Fig. 3), gametic phase disequilibrium reduced the advantage of genotypic selection over mass selection in early generations. In generation 15, however, the relative advantage of mass selection over genotypic selection was unaffected by the presence of gametic phase disequilibrium.

The continuous lines in Fig. 4 show the amount of gametic phase disequilibrium generated between the major gene and polygenes under the different types of selection ( $\bar{A}_{b,t} - \bar{A}_{B,t}$ ). For both mass selection and genotypic selection, the amount of gametic phase disequilibrium increased during the first three generations and then reached a plateau, as expected. Gametic phase disequilibrium was more than twice as large for genotypic selection as for mass selection.

Differences in responses between mass and genotypic selection are caused by emphasis put on the major gene versus polygenes over the course of selection. In terms of the index used for selection

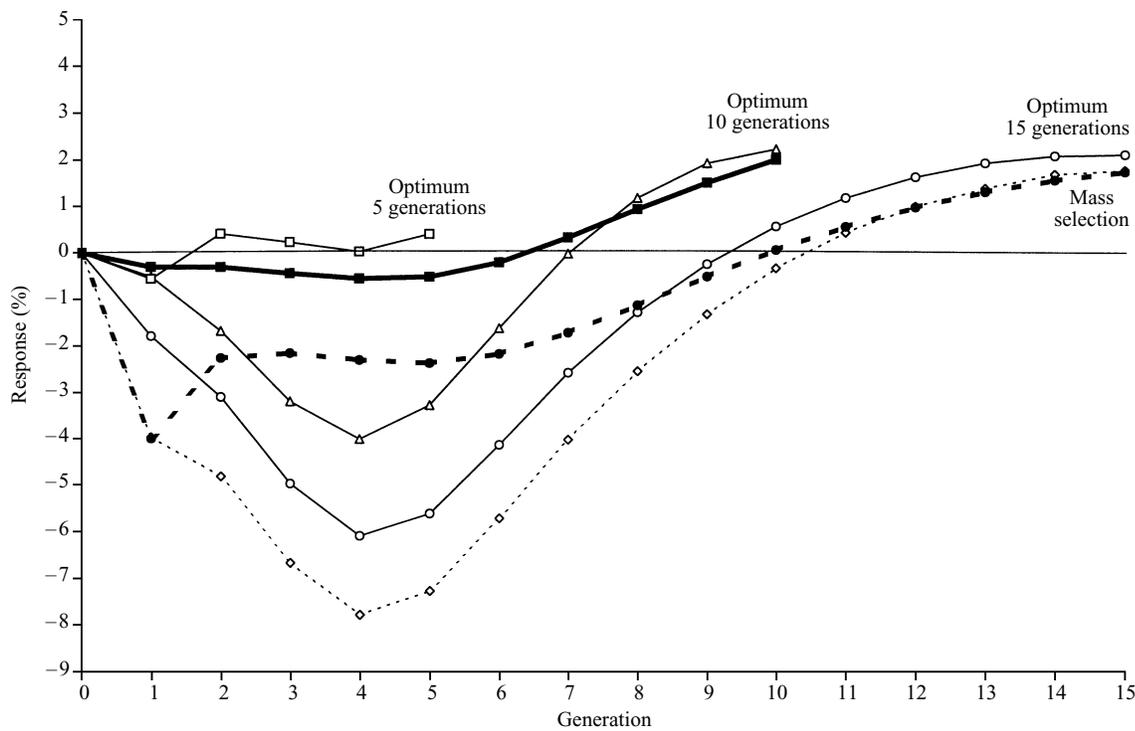


Fig. 3. Cumulative total response (major gene plus polygenes) to mass selection (broken lines) and optimum selection (continuous lines) for a quantitative trait with a segregating major gene, as a percentage deviation of response for genotypic selection, and for models without (thin lines) and with (thick lines) consideration of gametic phase disequilibrium between the major gene and polygenes. Under optimum selection, results are shown for maximization of response over 5, 10, or 15 generations for the model without gametic phase disequilibrium and over 10 generations for the model with gametic phase disequilibrium.

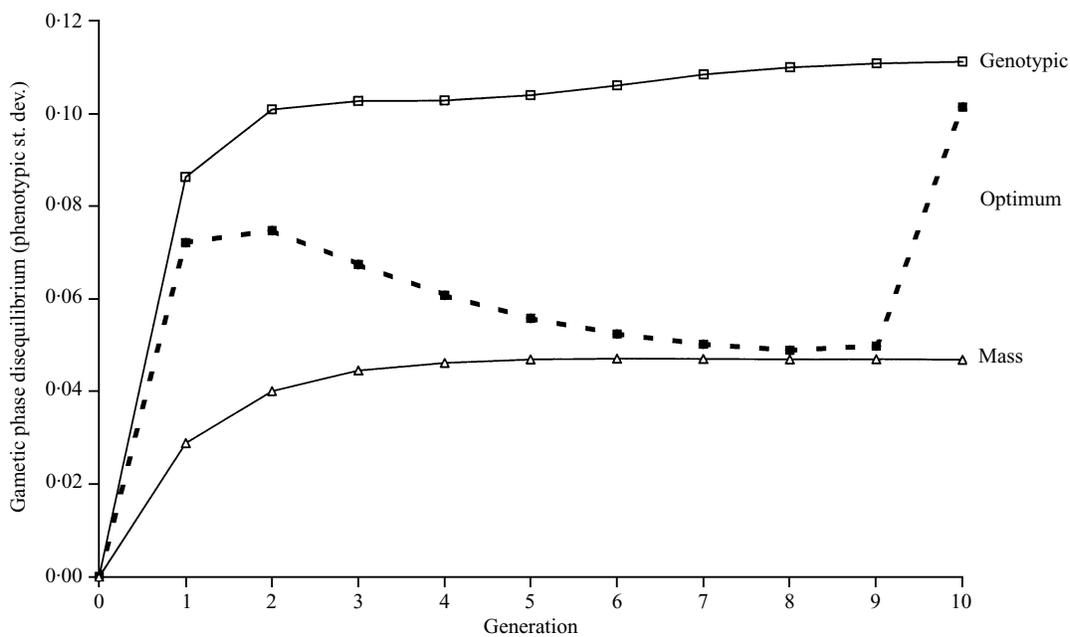


Fig. 4. Gametic phase disequilibrium between the major gene and polygenes for mass selection, genotypic selection and optimum selection. Gametic phase disequilibrium in generation  $t$  is defined as the average polygenic value (in phenotypic standard deviation units) of gametes which form generation  $t$  and contain the undesirable major gene allele minus the average polygenic value of gametes that contain the favourable major gene allele.

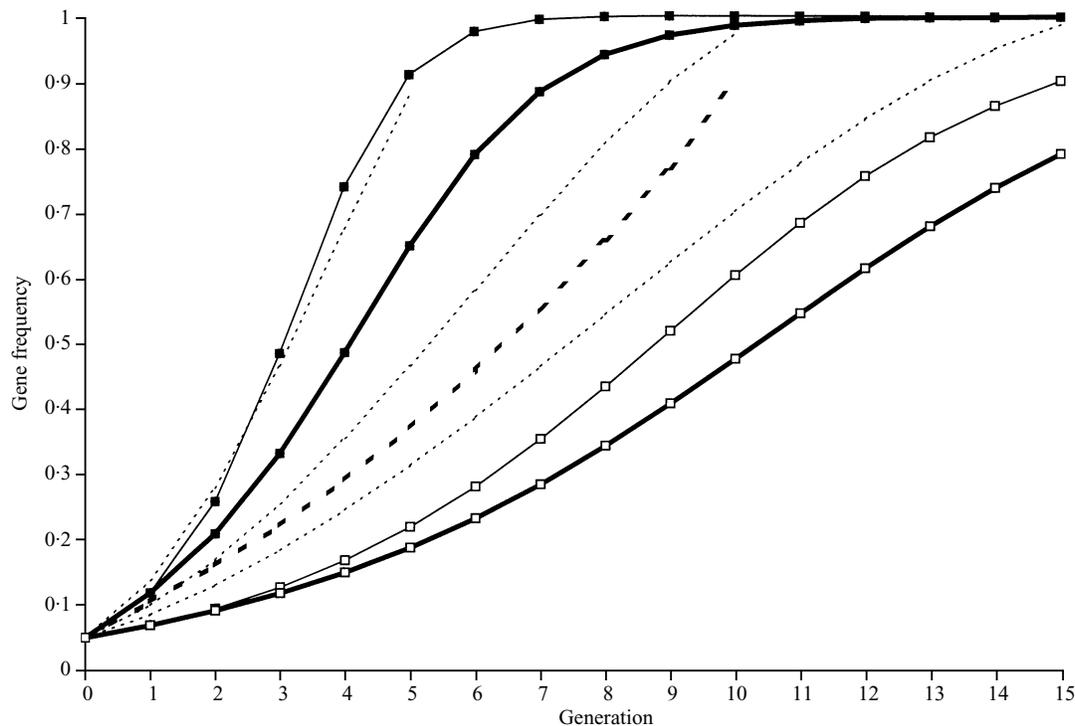


Fig. 5. Changes in frequency of the favourable allele of a major gene for a quantitative trait in response to mass selection (open squares), genotypic selection (filled squares) and optimum selection (broken lines), for a model without (thin lines) and with (thick lines) consideration of gametic phase disequilibrium between the major gene and polygenes. For optimum selection, results are shown for maximization of response over 5, 10 or 15 generations for the model without gametic phase disequilibrium and over 10 generations for the model with gametic phase disequilibrium.

(equation (1)), weights on major gene effects were 0.3 ( $=h^2$ ) for mass selection and 1.0 for genotypic selection. Weights were unaffected by gametic phase disequilibrium, assuming that actual differences between major genotypes [i.e.  $a - (\bar{A}_{b,t} - \bar{A}_{B,t})$ ] can be measured without error in each generation. Greater emphasis on the major gene resulted in greater changes in frequency of the major gene for genotypic compared with mass selection, which is illustrated in Fig. 5 (continuous lines). Presence of gametic phase disequilibrium reduced rates of increase in gene frequency for both genotypic selection and mass selection (Fig. 5). Gametic phase disequilibrium reduced the magnitude of effects associated with the major gene in the population from  $a$  to  $a - (\bar{A}_{b,t} - \bar{A}_{B,t})$  and, therefore, reduced effective selection pressure on the major gene.

Relative rates of improvement in polygenic breeding values are presented in Fig. 6 (broken lines). Fig. 6 illustrates that mass selection put more selection pressure on polygenic effects and, as a result, achieved higher rates of response in polygenic effects.

#### (ii) Optimum selection

For optimum selection, total cumulative response relative to genotypic selection is illustrated in Fig. 3

(continuous lines), changes in frequency of the major gene in Fig. 5 (broken lines) and cumulative polygenic response relative to polygenic response with genotypic selection in Fig. 6 (continuous lines). Results are presented for three planning horizons (maximization of cumulative response in generations 5, 10 and 15) for the model without gametic phase disequilibrium and for one planning horizon (10 generations) for the model with gametic phase disequilibrium. The broken line in Fig. 4 shows the extent of gametic phase disequilibrium generated for the latter.

In all cases, optimum selection achieved greater cumulative total response at the end of the planning horizon than genotypic or mass selection (Fig. 3, continuous lines). Without gametic phase disequilibrium, optimum selection resulted in 0.4%, 2.2% and 2.1% greater cumulative response than genotypic selection for planning horizons of 5, 10 and 15 generations, respectively. Comparing relative cumulative response at the end of the planning horizon for optimum selection (continuous lines in Fig. 3) with relative cumulative response over corresponding planning horizons for mass selection (broken line in Fig. 3), optimum selection resulted in 8.3%, 2.6% and 0.3% greater cumulative responses than mass selection for planning horizons of 5, 10 and 15 generations, respectively. Therefore, for the situation considered

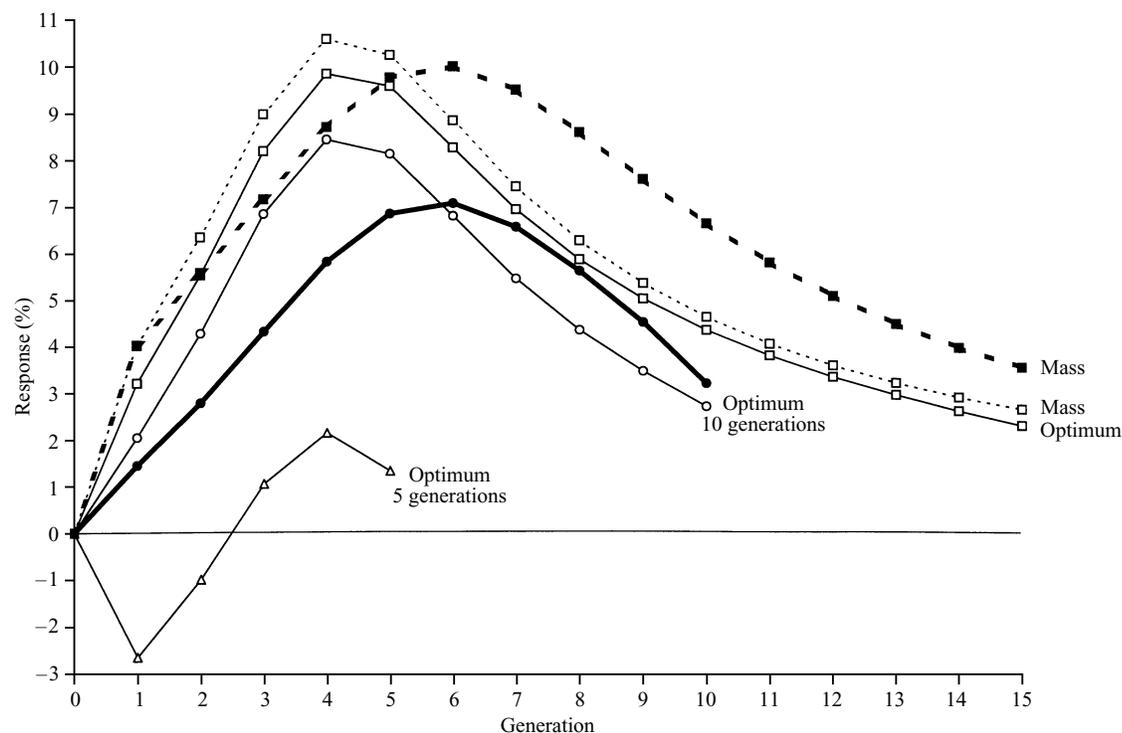


Fig. 6. Cumulative polygenic response to mass selection (broken lines) and optimum selection (continuous lines) for a quantitative trait with a segregating major gene, as a percentage deviation of polygenic response for genotypic selection, and for models without (thin lines) and with (thick lines) consideration of gametic phase disequilibrium between the major gene and polygenes. Under optimum selection, results are shown for maximization of response over 5, 10 or 15 generations for the model without gametic phase disequilibrium and over 10 generations for the model with gametic phase disequilibrium.

here, genotypic selection was close to optimum for short planning horizons but mass selection was closer to optimum for longer planning horizons.

Extra cumulative response for optimum selection relative to genotypic or mass selection at the end of a planning horizon of 10 generations was little affected by the presence of gametic phase disequilibrium between the major gene and polygenes (Fig. 3, continuous lines). Without gametic phase disequilibrium, cumulative response in initial generations was substantially less for optimum selection over 10 or 15 generations than cumulative response for genotypic selection (Fig. 3). Cumulative response relative to genotypic selection was reduced less in initial generations for optimum selection over 10 generations with than without gametic phase disequilibrium (Fig. 3).

In every generation, cumulative response in polygenic breeding values to optimum selection was intermediate to polygenic response for mass and genotypic selection (Fig. 6, continuous lines). Exceptions were the first 2 generations for a planning horizon of 5 generations, for which optimum selection achieved less polygenic response than genotypic selection. In generation 10, cumulative polygenic response was 2.5% and 3.0% greater for optimum

selection than for genotypic selection, without and with gametic phase disequilibrium, respectively. In contrast, changes in frequency of the major gene were greater for genotypic selection than for optimum selection (Fig. 5, broken lines). For optimum selection, changes in gene frequency were intermediate to those for mass selection and genotypic selection, depending on length of the planning horizon: with a short planning horizon (5 generations), changes in gene frequency for optimum selection were close to those observed for genotypic selection; for longer planning horizons (15 generations), changes in gene frequency for optimum selection tended to be more similar to those found for mass selection.

Differences in responses to optimum selection relative to mass and genotypic selection are caused by weights put on major gene effects in the selection index (equation (1)). Weights for optimum selection are presented in Fig. 7. Note that weights for genotypic and mass selection were constant over generations at 1.0 and  $h^2$  (= 0.3), respectively. Weights for optimum selection changed over generations, decreasing during the first half of the planning horizon and increasing during the second half. Optimum weights in the last generation of selection were always equal to weights under genotypic selection. This is explained by the

fact that the objective in the last generation of selection is to maximize response in the next generation, given the cumulative gain obtained up to that point in time. Consequences for subsequent generations are no longer considered. Genotypic selection (index weight = 1) maximized response from one generation to the next, at least for the additive genetic models considered here.

Without gametic phase disequilibrium,  $b_{1t} = b_{3t}$  for optimum selection (see Appendix) and, therefore, only one line is shown in Fig. 7 for each planning horizon (continuous lines). With gametic phase disequilibrium, however, weights  $b_{3t}$  were greater than weights  $b_{1t}$  for all generations except the last (Fig. 7, broken lines). In the last generation,  $b_{1t} = b_{3t}$  and equal to the weights under genotypic selection (= 1), as expected. The fact that  $b_{3t} > b_{1t}$  illustrates that, when gametic phase disequilibrium was accounted for, the optimum index put more emphasis on selection against the undesired bb genotype ( $b_{3t}$ ) than on selection in favour of BB ( $b_{1t}$ ).

In general, index weights for optimum selection were intermediate to those for mass selection and genotypic selection (Fig. 7). With a planning horizon of 5 generations, however, emphasis on the major gene was greater in the first generation for optimum selection than for genotypic selection.

Selection index weights quantify the weight that is put on genetic values for the major gene relative to polygenic EBV when computing the selection criterion. Index weights do not, however, quantify the effective selection pressure that is put on the major gene relative to polygenes, which also depends on the amount of variation that is present in the population. For example, although genotypic selection maintains a constant weight of one on the major gene in selection index (1) over generations, the effective selection pressure on the major gene will be lower when the frequency of the major gene is close to the extremes (0 or 1) because most animals will be of the same genotype. Therefore, at low or high frequency of the major gene, a weight of 1 on the major genotype value, as in genotypic selection, will have less impact on selection for polygenic breeding values than when the major gene is at intermediate frequency.

To better quantify effective selection pressure on the major gene and polygenes, measures of achieved selection intensity were considered. Achieved selection intensity for the major gene (or polygenes) was computed as the ratio of response achieved for the major gene (polygenes) over the square root of variance contributed to the selection criterion by the major gene (polygenes) ( $= 2(p_{t+1} - p_t) / [2p_t(1 - p_t)]^{0.5}$  for the major gene and  $= (\bar{A}_{t+1} - \bar{A}_t) / \sigma$  for polygenes). These derivations stem from the fact that expected response to selection is equal to intensity times the standard deviation of EBV ( $= \text{intensity} \times \text{accuracy} \times$

genetic standard deviation) (Falconer & Mackay, 1996). Results are in Fig. 8.

For genotypic selection, intensity on the major gene varied over generations (Fig. 8), although the index weight on the major gene remained constant (= 1). Changes in intensity were due to changes in variance contributed by the major gene as gene frequency changed over generations. Correspondingly, intensity achieved for polygenes also changed over generations, but in an opposite direction. Interestingly, intensity on the major gene and polygenes was constant over generations under optimum selection, at least under the model without gametic phase disequilibrium. Intensity varied to some degree when gametic phase disequilibrium was considered but much less than with genotypic selection.

#### 4. Discussion and conclusions

The main objective of this paper was to develop a theoretical framework for methods to optimize response to selection over multiple generations or multiple stages of selection when molecular genetic information is available. Consideration of selection over more than one generation or stage becomes important when population parameters (e.g. heritability, genetic variance, gene frequencies) change over generations as a result of selection or other factors. This is the case for most genetic systems. In most of these cases, traditional methods of selection for quantitative traits on breeding values that are estimated based on selection index or BLUP maximize response from one generation to the next. The changes in parameters that result from this selection have, however, consequences for responses to selection that can be achieved in subsequent generations. Situations in which selection on BLUP of breeding values does not maximize response over more than one generation were discussed by Woolliams (1990).

Dekkers *et al.* (1995) proposed use of optimal control theory as a method to formulate and solve multiple-generation selection problems and applied this method to optimize selection over multiple generations with non-linear profit functions. Several other potential applications of optimal control theory to multiple-generation selection problems in animal breeding were discussed, including optimization of selection over multiple generations with gametic phase equilibrium, overlapping generations, and optimization of selection with inbreeding. The specific application of optimal control to multiple-generation selection problems addressed in the present study was maximization of the longer-term response to selection on a quantitative trait when information on a single gene is available. Results illustrate that selection based on information from identified genes can be

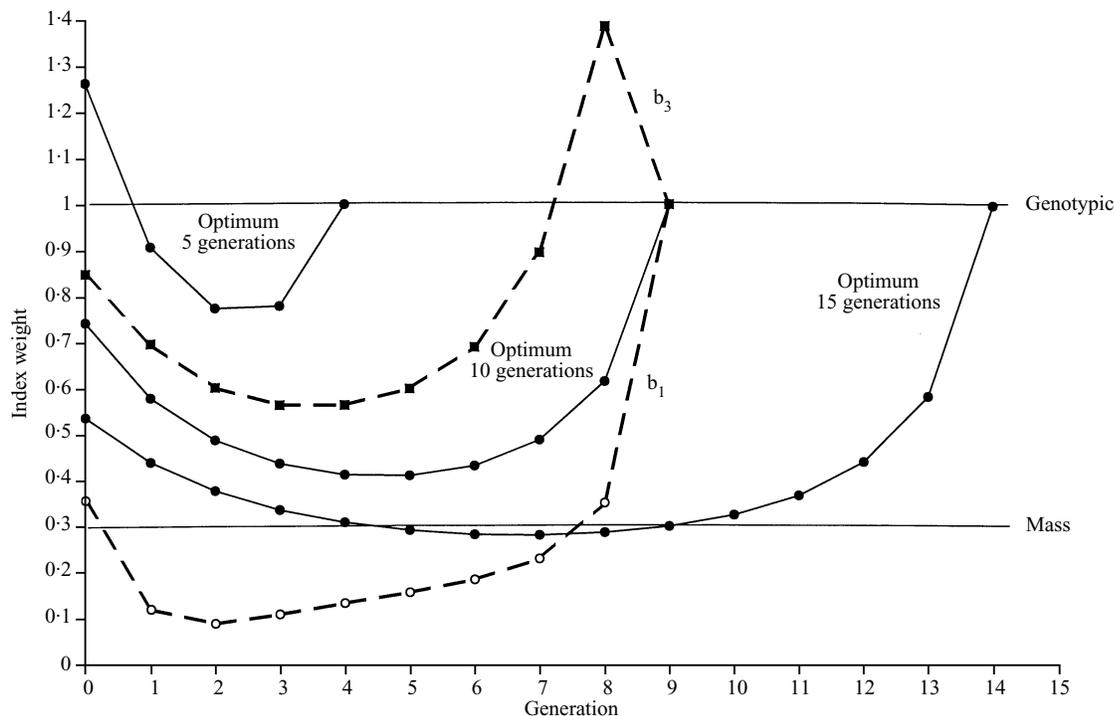


Fig. 7. Index weights on the major gene for optimum selection with maximization of cumulative response over 5, 10 and 15 generations for a model without consideration of gametic phase disequilibrium between the major gene and polygenes (continuous lines) and with maximization of cumulative response over 10 generations for a model with consideration of gametic phase disequilibrium (broken lines). For the model without consideration of gametic phase disequilibrium, index weights are equal for alternative major genotypes. For the model with gametic phase disequilibrium, the index weight on the homozygous favourable genotype ( $b_1$ ) differs from the weight on the homozygous unfavourable genotype ( $b_3$ ). For comparison, index weights are also shown for genotypic selection (= 1) and mass selection (= heritability of the trait).

optimized and that optimal control theory provides a useful framework to formulate and optimize such selection systems.

One of the main conclusions of this paper is that a build-up of gametic phase disequilibrium between the major gene and polygenes is not the main reason why genotypic selection results in less than optimal responses to selection in the medium and long term. Instead, suboptimality of genotypic selection is mainly caused by the fact that selection pressure and response for polygenes changes over generations with genotypic selection. This unequal selection pressure on polygenes is caused by changes in frequency for the major gene, which changes the amount of variance that is contributed by the major gene and, therefore, changes the selection pressure that is devoted to the major gene versus polygenes. Optimum selection, i.e. selection that maximizes cumulative response over a planning horizon of multiple generations, resulted in constant selection pressure on the two components that contribute to total response, i.e. the major gene and polygenes (Fig. 8), at least under the model without gametic phase disequilibrium. In other words, optimum selection balanced reductions in polygenic response over generations, in contrast to genotypic selection. Equalization of reductions in polygenic

response over generations resulted in minimum cumulative reductions in polygenic response, while maximizing total response to selection. Optimum selection achieves this by taking into account the effect of current selection decisions on future changes in frequency for the major gene and the effect of major gene frequency on variance and response contributed by the major gene. In a related study, Luo *et al.* (1997) also observed that gene frequencies at an unknown QTL and at a linked genetic marker had important and non-linear effects on responses to marker-assisted selection. This relationship between frequency and response at the major gene, and its consequences for selection pressure on polygenes, seems to be an important factor that forms the basis for the difference between genotypic and optimum selection in the present study.

It is interesting to note that the result that selection pressure on components that contribute to response to selection is constant under optimum selection has similarities with results obtained by Dekkers *et al.* (1995) for optimum selection on non-linear profit functions, although the scenarios considered are distinctly different: for selection on non-linear profit functions, the selection index that maximized cumulative response over a given planning horizon also

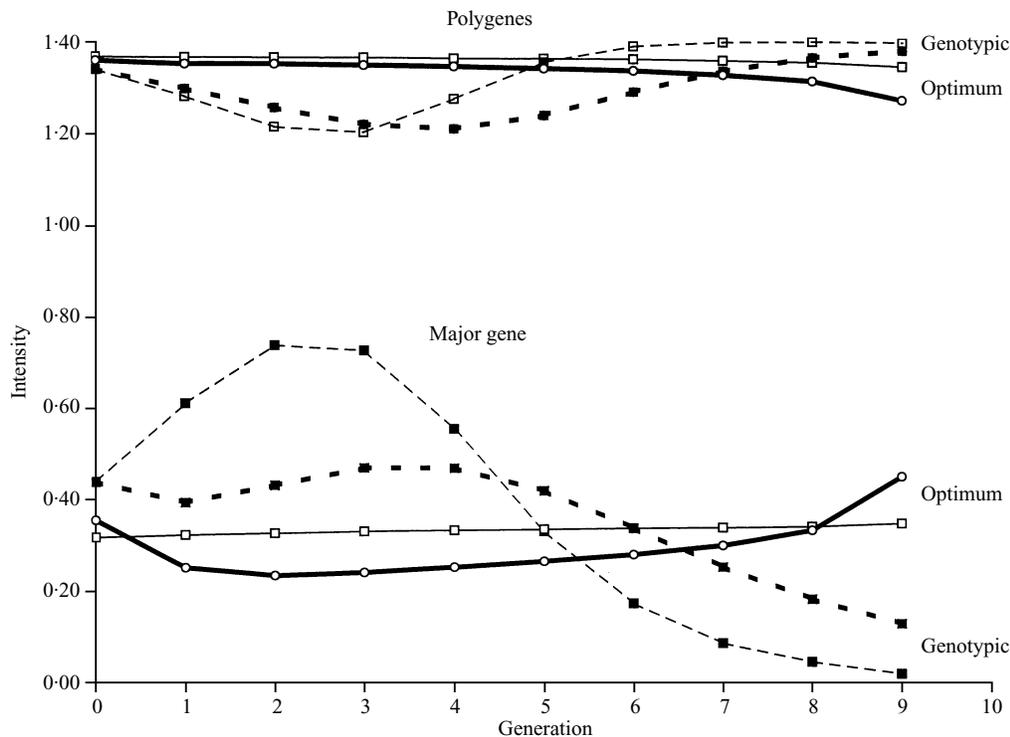


Fig. 8. Achieved selection intensities on polygenes and the major gene with genotypic selection (broken lines) and optimum selection over 10 generations (continuous lines) for a model without (thin lines) or with (thick lines) consideration of gametic phase disequilibrium between the major gene and polygenes. For polygenes, selection intensity in generation  $t$  was computed as  $(\bar{A}_{t+1} - \bar{A}_t)/\sigma$ , where  $\bar{A}_t$  is equal to the average polygenic value in generation  $t$  and  $\sigma$  is the standard deviation of estimates of polygenic breeding values. Selection intensity achieved for the major gene was computed as  $(\text{major gene response})/(\text{population variance of major gene effects})^{0.5} = 2(p_{t+1} - p_t)/[2p_t(1 - p_t)]^{0.5}$ , where  $p_t$  is the frequency of the major gene in generation  $t$ .

resulted in constant achieved selection intensity (and response) for each of the traits that contributed to the selection goal.

For the model with gametic phase equilibrium, achieved selection intensities on the major gene and polygenes were not entirely constant over generations under optimum selection (Fig. 8). This suggests that additional factors play a role. The relatively small deviations from constant selection intensities illustrate, however, that the aforementioned relationship between frequency and variance at the major gene, and its consequences for selection pressure on polygenic breeding values, is an important factor that contributes to characteristics of the optimum strategy.

Models used in this paper were rather simplistic with regard to genetics (no dominance, gametic phase equilibrium among polygenes, and no linkage between the major gene and polygenes), structure of the breeding programme (discrete generations, equal selection in both sexes, and infinite population size), and objective function (maximization of cumulative response over a planning horizon rather than, for example, maximization of cumulative discounted responses to selection; Dekkers *et al.*, 1995). Methods and examples presented in this paper do, however,

provide important insight into the process of longer-term selection on quantitative traits with information on identified genes and into methods that can be used to optimize such selection strategies. In principle, methods developed herein can be extended to more complex situations.

The main limitation of the genetic model used here in relation to accepted genetic models is the assumption of constant polygenic variance, which resulted from the assumption of gametic phase equilibrium among polygenes. Selection reduces polygenic variance to an extent which depends on selection intensity and selection accuracy (Bulmer, 1980). For quantitative traits that are affected by a major gene and polygenes, parents with two copies of the unfavourable major gene allele (bb) are selected with greater intensity for polygenic effects than parents with two copies of the favourable major gene allele (BB) (Fig. 2). Hence, polygenic variance will be reduced to a greater degree for bb parents than for BB parents. Accounting for this factor will favour selection of BB animals compared with the present model. The impact on optimum selection is expected to be that the difference in weights  $b_{1t}$  and  $b_{3t}$  will be less than was observed for the current model (Fig. 7).

This would reduce the difference between optimum and genotypic selection. In contrast, however, the magnitude of the effect of the major gene will increase relative to the effect of polygenes when the standard deviation of polygenic EBV is reduced as a result of selection. A larger major gene effect relative to polygenic EBV will increase the difference between index weights as well as the benefit of optimum selection versus genotypic selection (results not shown). As a result of these two opposing effects, general trends and characteristics of the optimum strategies may be similar to results discussed here for models that include gametic phase disequilibrium among polygenes.

Including gametic phase disequilibrium among polygenes will further erode the main phenomenon that was observed here for the model without gametic phase disequilibrium between the major gene and polygenes, i.e. that optimum selection results in constant selection pressure and response to selection for polygenes. With gametic phase disequilibrium among polygenes, current selection decisions affect not only future genetic variance contributed by the major gene but also genetic variance due to polygenes. A related assumption of the present model, however, was that the base population was unselected and in gametic phase equilibrium. If marker-assisted selection is implemented in an ongoing selection programme that is based on EBV estimated from phenotype, a degree of gametic phase disequilibrium will already be established, both among polygenes and between the major gene and polygenes. As a result, further changes in polygenic variance and in gametic phase disequilibrium between the major gene and polygenes will be reduced. Results would, therefore, be expected to be more similar to what was observed here for the model that did not include gametic phase disequilibrium between the major gene and polygenes, nor gametic phase disequilibrium among polygenes. Further development of models to accommodate these factors is, however, needed. The theoretical framework developed here can serve as the basis of such further developments.

In general, benefits of optimum selection over genotypic selection were small for the example studied here (Fig. 3). It is clear that benefits of optimum selection depend on frequency and size of the major gene and on length of the planning horizon that is considered in the objective. In this study, only a few examples were considered and the objective involved maximization of cumulative response at the end of a planning horizon. With a starting frequency for the major gene of 5%, the advantage of optimum over genotypic selection was small (0.4%) for a planning horizon of 5 generations and increased to 2.2% for planning horizons of 10 and 15 generations (Fig. 2). Genotypic selection is equivalent to selection on

BLUP of total breeding values (Kennedy *et al.*, 1992) and maximizes response in the next generation for major genes with additive effects. Genotypic selection can, therefore, be considered a short-term selection criterion. It must also be noted that 5 generations is considered long for most livestock species. Recent studies (Dekkers, in preparation) have, however, shown that genotypic selection does not maximize short-term response if the major gene exhibits dominance.

Benefits of optimum selection over genotypic selection also depend on starting frequency of the major gene. For example, with a starting frequency of 25% (results not shown), optimum selection for a planning horizon of 5 generations resulted in 1.9% greater response than genotypic selection (no gametic phase disequilibrium). This compared with only 0.4% greater response when the starting frequency was 5% (Fig. 3).

In the problem addressed here, the identified gene had a direct effect on the quantitative trait of interest. With some modification, the same method can, however, also be applied to selection against an undesirable single gene that has no effect on the quantitative trait. Examples are the halothane gene in swine (Eikelenboom & Minkema, 1974; Fuji *et al.*, 1991) and the BLAD gene in dairy cattle (Shuster *et al.*, 1992). In these cases, the objective could be to eliminate the gene without sacrificing large responses to selection for a quantitative trait of interest (e.g. growth in the case of swine or milk production in the case of dairy cattle). For a gene with additive effects, the goal of selection over  $T$  generations can then be formulated from an economic perspective in a manner similar to (2) as:  $\bar{G}_T = a(2p_T - 1) + v\bar{A}_T$ , where  $a$  and  $v$  quantify the relative economic importance of the undesirable gene and the quantitative trait, respectively. Selection is then on an index of a form similar to that used in the current paper (equation (1)) and optimum selection procedures can be derived based on the methods developed here. If the single gene has pleiotropic effects on the quantitative trait of interest, as appears to be the case for the halothane gene in swine, this can be accounted for by modifying relative economic values  $a$  and  $v$ . Methods can also be extended to situations with non-additive gene effects and to situations in which the objective is to eliminate the undesirable gene within a certain number of generations while minimizing loss of response for the quantitative trait. Extension to selection on markers linked to QTL is less straightforward because recombination between the genetic marker and major gene must be taken into account.

For the example investigated in this paper, benefits of selection procedures that used genotypic information (i.e. genotypic or optimum selection) were limited to selection procedures that used phenotypic

information only (i.e. mass selection in the present example), in particular in the longer term (Fig. 3). Although benefits were greater in the short term, gains were less than 4% when gametic phase disequilibrium was considered (Fig. 7). Similar results were found by Ruane & Colleau (1995) for marker-assisted selection. Several studies have, however, shown, that greater gains can be expected from use of information from single genes or genetic markers if traits have low heritability (e.g. Smith & Simpson, 1986; Ruane & Colleau, 1996) or are sex-limited (e.g. Van der Beek & Van Arendonk, 1994; Ruane & Colleau, 1996) and at stages of selection for which limited information is available on, in particular, the Mendelian sampling component that is received by the animal from its parents, i.e. prior to availability of own phenotype or progeny records (e.g. Kashi *et al.*, 1990; Meuwissen & Van Arendonk, 1992). It is expected that in these cases the advantage of optimum over genotypic selection will also be greater than observed for the example in the current study. In addition, most studies on marker-assisted selection have evaluated selection on genes with additive effects. With dominance at the major gene, benefits of optimum over genotypic selection are expected to be substantially greater, even in the short term (Dekkers, in preparation).

**Appendix. Maximization of cumulative response to selection over a planning horizon of T generations**

(i) *No gametic phase disequilibrium between the major gene and polygenes*

After incorporation of constraints using Lagrange multipliers, maximization of cumulative response to selection over *T* generations (equations (5)) amounts to maximization of the following function (from equations (6), (7) and (8)):

$$\text{Max}_{f_{mt}} \{L\} \text{ given } \bar{A}_0, p_0 \text{ and } Q \tag{A 1}$$

with

$$L = \sum_{t=0}^{T-1} \{H_t - \lambda_t p_t - \gamma_t \bar{A}_t\} - \lambda_T p_T + \lambda_0 p_0 - \gamma_T \bar{A}_T + \gamma_0 \bar{A}_0 + a(2p_T - 1) + \bar{A}_T, \tag{A 2}$$

where *H<sub>t</sub>* is the Hamiltonian (Lewis, 1986), which is defined for *t* = 0 to *T* - 1 and is equal to:

$$H_t = \frac{\lambda_{t+1}}{Q} \{f_{1t} p_t^2 + f_{2t} p_t(1 - p_t)\} + \gamma_{t+1} \times \bar{A}_t + \frac{\sigma}{Q} [p_t^2 z_{1t} + 2p_t(1 - p_t) z_{2t} + (1 - p_t)^2 z_{3t}] + \epsilon_t \{Q - f_{1t} p_t^2 - f_{2t} 2p_t(1 - p_t) - f_{3t}(1 - p_t)^2\}. \tag{A 3}$$

Equating first partial derivatives of *L* to zero results in the following set of equations for *t* = 0 to *T* - 1:

$$\frac{\delta L}{\delta f_{1t}} = \frac{\delta H_t}{\delta f_{1t}} = p_1^2 \frac{\lambda_{t+1}}{Q} + \frac{\sigma \gamma_{t+1} \delta z_{1t}}{Q \delta f_{1t}} - \epsilon_t = 0, \tag{A 4a}$$

$$\frac{\delta L}{\delta f_{2t}} = \frac{\delta H_t}{\delta f_{2t}} = p_t(1 - p_t) \frac{\lambda_{t+1}}{Q} + 2 \frac{\sigma \gamma_{t+1} \delta z_{2t}}{Q \delta f_{2t}} - 2\epsilon_t = 0, \tag{A 4b}$$

$$\frac{\delta L}{\delta f_{3t}} = \frac{\delta H_t}{\delta f_{3t}} = (1 - p_t)^2 \frac{\sigma \gamma_{t+1} \delta z_{3t}}{Q \delta f_{3t}} - \epsilon_t = 0, \tag{A 4c}$$

$$\frac{\delta L}{\delta p_t} = \frac{\delta H_t}{\delta p_t} - \lambda_t = \frac{\lambda_{t+1}}{Q} \{2f_{1t} p_t + f_{2t}(1 - 2p_t)\} + 2 \frac{\sigma \gamma_{t+1}}{Q} \{p_t z_{1t} + (1 - 2p_t) z_{2t} - (1 - p_t) z_{3t}\} - 2\epsilon_t \{p_t f_{1t} + (1 - 2p_t) f_{2t} + (1 - p_t) f_{3t}\} - \lambda_t = 0, \tag{A 4d}$$

$$\frac{\delta L}{\delta A_t} = \frac{\delta H_t}{\delta A_t} - \gamma_t = \gamma_{t+1} - \gamma_t = 0, \tag{A 4e}$$

$$\frac{\delta L}{\delta \lambda_{t+1}} = \frac{\delta H_t}{\delta \lambda_{t+1}} - p_{t+1} = \frac{1}{Q} \{f_{1t} p_t^2 + f_{2t} p_t(1 - p_t)\} - p_{t+1} = 0, \tag{A 4f}$$

$$\frac{\delta L}{\delta \gamma_{t+1}} = \frac{\delta H_t}{\delta \gamma_{t+1}} - \bar{A}_{t+1} = \bar{A}_t + \frac{\sigma}{Q} \{p_t^2 z_{1t} + 2p_t(1 - p_t) z_{2t} + (1 - p_t)^2 z_{3t}\} - \bar{A}_{t+1} = 0, \tag{A 4g}$$

$$\frac{\delta L}{\delta \epsilon_t} = \frac{\delta H_t}{\delta \epsilon_t} = Q - f_{1t} p_t^2 - f_{2t} 2p_t(1 - p_t) - f_{3t}(1 - p_t)^2 = 0; \tag{A 4h}$$

and for *t* = *T*:

$$\frac{\delta L}{\delta p_T} = -\lambda_T + 2a = 0, \tag{A 4i}$$

$$\frac{\delta L}{\delta A_t} = -\gamma_T + 1 = 0. \tag{A 4j}$$

Equation (A 4e) results in  $\gamma_{t+1} = \gamma_t$  for *t* = 0 to *T* - 1, which along with  $\gamma_T = 1$  (from (A 4j)), results in  $\gamma_t = 1$  for all *t*. Variable  $\gamma$ , represents the shadow value for  $\bar{A}_t$ .

Using  $\delta z_{mt} / \delta f_{mt} = x_{mt}$ , which is based on properties of the standard Normal distribution, (A 4a), (A 4b) and (A 4c), along with (A 4h), can be used to solve for optimum control variables in generation *t* (*f<sub>mt</sub>*), given the Lagrange multipliers for *t* + 1 ( $\lambda_{t+1}$ ), as described below. From (A 4a):

$$\epsilon_t = \frac{1}{Q} (\lambda_{t+1} + \sigma x_{1t}). \tag{A 5a}$$

From (A 4b):

$$\epsilon_t = \frac{1}{Q} (\frac{1}{2}\lambda_{t+1} + \sigma x_{2t}). \tag{A 5b}$$

From (A 4c):

$$\epsilon_t = \frac{1}{Q} \sigma x_{3t}. \tag{A 5c}$$

Combining (A 5a) and (A 5b) results in:

$$x_{1t} = x_{2t} - \frac{1}{2\sigma} \lambda_{t+1}. \tag{A 6a}$$

Combining (A 5b) and (A 5c) results in:

$$x_{3t} = x_{2t} + \frac{1}{2\sigma} \lambda_{t+1}. \tag{A 6b}$$

Equations (A 6a) and (A 6b), which set standardized truncation points given  $x_{2t}$ , along with (A 4b), which returns constraints (5a) and sets the overall selected fraction, can be used to derive the optimum truncation points and fractions selected in generation  $t$ , given  $p_t$  and  $\lambda_{t+1}$ . Iterative procedures of Ducrocq & Quaas (1988) can be used for this purpose.

Note that combining (A 6a) and (A 6b) results in:

$$x_{2t} - x_{1t} = x_{3t} - x_{2t} = \frac{1}{2\sigma} \lambda_{t+1}, \tag{A 7}$$

which implies that, in every generation  $t$ , optimum standardized truncation points are equidistant for the three genotype classes. This is a result of the modelling of linkage phase equilibrium.

Solving (A 6a) and (A 6b) for generation  $t$  depend on knowing  $\lambda_{t+1}$  and  $p_t$ . Using (A 4d), the following backward recursive equation can be derived for  $\lambda_t$ :

$$\begin{aligned} \lambda_t = \frac{\lambda_{t+1}}{Q} \{2f_{1t}p_t + f_{2t}(1-2p_t)\} \\ + \frac{\sigma}{Q} \{2p_t z_{1t} + 2(1-2p_t)z_{2t} - 2(1-p_t)z_{3t}\} \\ + \{-2\epsilon_t p_t f_{1t} - 2\epsilon_t(1-2p_t)f_{2t} - 2\epsilon_t(1-p_t)f_{3t}\}, \end{aligned} \tag{A 8}$$

with a starting point,  $\lambda_T = 2a$ , which is obtained from (A 4i). Substituting (A 5a), (A 5b) and (A 5c) in respectively the first, second and third terms of (A 8) that contain  $\epsilon_t$ , simplifies (A 8) into:

$$\begin{aligned} \lambda_t = \frac{2\sigma}{Q} \{p_t(z_{1t} - f_{1t}x_{1t}) + (1-2p_t)(z_{2t} - f_{2t}x_{2t}) \\ + (1-p_t)(f_{3t}x_{3t} - z_{3t})\}. \end{aligned} \tag{A 9}$$

Lagrange multipliers  $\lambda_t$  can be removed from the solution procedure by substituting (A 6b) into (A 9), which results in the following backward recursive equation for  $(x_{3t} - x_{2t})$ :

$$\begin{aligned} (x_{3,t-1} - x_{2,t-1}) = \frac{1}{Q} \{p_t f_{1t} (i_{1t} - x_{1t}) \\ + (1-2p_t)f_{2t}(i_{2t} - x_{2t}) + (p_t - 1)f_{3t}(i_{3t} - x_{3t})\}, \end{aligned} \tag{A 10}$$

which has as starting point (from  $\lambda_T = -2a$ ):  $(x_{3T} - x_{2T}) = a/\sigma$ . A recursive equation for  $p_t$  is obtained from (A 4f), which results in (5b).

Based on the above, optimal solutions must satisfy (A 4b), (A 4f), (A 6a), (A 6b) and (A 9), where the latter three sets of equations can be replaced by (A 10). Using these equations, the following iterative procedure can be used for finding the optimum:

1. Set  $(x_{3t} - x_{2t}) = (x_{2t} - x_{1t}) = a/\sigma$  for all  $t$ .
2. For  $t = 0$  and given  $p_0$ ,  $(x_{3,0} - x_{2,0})$  and  $(x_{2,0} - x_{1,0})$ , derive  $f_{m,0}$  that satisfy the constraint given by (A 4b) (overall fraction selected,  $Q$ ), using the truncation selection procedure of Ducrocq & Quaas (1988). Compute the associated values for  $z_{m,0}$  and  $i_{m,0}$ .
3. Compute  $p_{t+1}$  based on (A 4f).
4. Repeat steps 2 and 3 for  $t = 1$  to  $T-1$ .
5. Using (A 10), compute new values for  $(x_{3t} - x_{2t})$  for  $t = 0$  to  $T-2$  ( $(x_{3,T-1} - x_{2,T-1})$  remains equal to  $a/\sigma$ ), given solutions for  $p_t$  and  $f_{mt}$  obtained from step 4. A multiplicative relaxation factor may be required here, reducing changes in  $(x_{3t} - x_{2t})$  from one iteration to another, to allow convergence.
6. Repeat steps 2 to 5 until  $(x_{3t} - x_{2t})$  converges to a stable solution.
7. Given  $\bar{A}_0$  and the optimal solutions, compute  $\bar{A}_t$  for each generation  $t$  based on (A 4g) (or (5c)); compute  $G_T$  based on (2); compute  $b_{mt}$  based on (4a) and (4b).

Note that the starting values for this iterative procedure, which are set in step 1 [ $(x_{3t} - x_{2t}) = (x_{2t} - x_{1t}) = a/\sigma$ ], provide results for genotypic selection. Results for mass selection can be obtained from step 1 by setting  $(x_{3t} - x_{2t})$  equal to  $a/\sigma_p$ , where  $\sigma_p$  is the phenotypic standard deviation.

(ii) *With gametic phase disequilibrium between major gene and polygenes*

Similar to the situation without gametic phase disequilibrium, equations (12) can be reformulated to maximizing a function  $L$ , similar to (A 1), with the following Hamiltonian function, which is defined for  $t = 0$  for  $T-1$ :

$$\begin{aligned} H_t = \frac{k_{t+1}}{2Q} \{ [2f_{1t}p_t + f_{2t}(1-p_t)]W_{B,t} \\ + f_{2t}p_t W_{b,t} + p_t \sigma [p_t z_{1t} + (1-p_t)z_{2t}] \} \\ + \frac{\lambda_{t+1}}{2Q} \{ [2f_{3t}(1-p_t) + f_{2t}p_t] W_{b,t} + f_{2t}(1-p_t)W_{B,t} \\ + (1-p_t) \sigma [(1-p_t)z_{3t} + p_t z_{2t}] \} \\ + \frac{\gamma_{t+1}}{Q} \{ f_{1t}p_t^2 + f_{2t}p_t(1-p_t) \} \\ + \epsilon_t \{ Q - f_{1t}p_t^2 - f_{2t}2p_t(1-p_t) - f_{3t}(1-p_t)^2 \}. \end{aligned} \tag{A 11}$$

Taking partial derivatives of  $L$  with respect to control variables, state variables and Lagrange multipliers, and equating them to zero, results in the following set of necessary conditions for an optimum (using  $W_{B,t} = p_t \bar{A}_{B,t}$  and  $W_{b,t} = (1 - p_t) \bar{A}_{b,t}$  in some instances to simplify equations) for  $t = 0$  to  $T - 1$ :

$$\frac{\delta H_t}{\delta f_{1t}} = \frac{k_{t+1}}{2Q} (2\bar{A}_{Bt} + x_{1t} \sigma) + \frac{\gamma_{t+1}}{Q} - \epsilon_t = 0, \tag{A 12a}$$

$$\frac{\delta H_t}{\delta f_{2t}} = \frac{k_{t+1} + \lambda_{t+1}}{2Q} (\bar{A}_{Bt} + \bar{A}_{bt} + x_{2t} \sigma) + \frac{\gamma_{t+1}}{Q} - 2\epsilon_t = 0, \tag{A 12b}$$

$$\frac{\delta H_t}{\delta f_{3t}} = \frac{\lambda_{t+1}}{2Q} (2\bar{A}_{bt} + x_{3t} \sigma) - \epsilon_t = 0, \tag{A 12c}$$

$$\begin{aligned} \frac{\delta H_t}{\delta p_t} - \gamma_t &= \frac{k_{t+1}}{2Q} \{ (2f_{1t} - f_{2t}) W_{B,t} \\ &+ f_{2t} W_{b,t} + [2p_t z_{1t} + (1 - 2p_t) z_{2t}] \sigma \} \\ &+ \frac{\lambda_{t+1}}{2Q} \{ (-2f_{3t} + f_{2t}) W_{b,t} \\ &- f_{2t} W_{B,t} + [2(p_t - 1) z_{3t} \\ &+ (1 - 2p_t) z_{2t}] \sigma \} \\ &+ \frac{\gamma_{t+1}}{Q} \{ 2f_{1t} p_t + f_{2t} (1 - 2p_t) \} \\ &- \epsilon_t \{ 2f_{1t} p_t + 2f_{2t} (1 - 2p_t) + 2f_{3t} (p_t - 1) \} \\ &- \gamma_t = 0, \end{aligned} \tag{A 12d}$$

$$\begin{aligned} \frac{\delta H_t}{\delta W_{B,t}} - k_t &= \frac{k_{t+1}}{2Q} [2f_{1t} p_t + f_{2t} (1 - p_t)] \\ &+ \frac{\lambda_{t+1}}{2Q} [f_{2t} (1 - p_t)] - k_t = 0, \end{aligned} \tag{A 12e}$$

$$\begin{aligned} \frac{\delta H_t}{\delta W_{b,t}} - \lambda_t &= \frac{k_{t+1}}{2Q} [f_{2t} p_t] \\ &+ \frac{\lambda_{t+1}}{2Q} [2f_{3t} (1 - p_t) + f_{2t} p_t] - \lambda_t = 0, \end{aligned} \tag{A 12f}$$

$$\begin{aligned} \frac{\delta H_t}{\delta k_{t+1}} - W_{B,t+1} &= \frac{1}{2Q} \{ [2f_{1t} p_t + f_{2t} (1 - p_t)] W_{B,t} + f_{2t} p_t W_{b,t} \\ &+ [p_t^2 z_{1t} + p_t (1 - p_t) z_{2t}] \sigma \} - W_{B,t+1} = 0, \end{aligned} \tag{A 12g}$$

$$\begin{aligned} \frac{\delta H_t}{\delta \lambda_{t+1}} - W_{b,t+1} &= \frac{1}{2Q} \{ [2f_{3t} (1 - p_t) + f_{2t} p_t] W_{b,t} \\ &+ f_{2t} (1 - p_t) W_{B,t} + [(1 - p_t)^2 z_{3t} \\ &+ p_t (1 - p_t) z_{2t}] \sigma \} - W_{b,t+1} = 0, \end{aligned} \tag{A 12h}$$

$$\frac{\delta H_t}{\delta \gamma_{t+1}} - p_{t+1} = \frac{1}{Q} [f_{1t} p_t^2 + f_{2t} p_t (1 - p_t)] - p_{t+1} = 0, \tag{A 12i}$$

$$\frac{\delta H_t}{\delta \epsilon_t} = Q - f_{1t} p_t^2 - f_{2t} 2p_t (1 - p_t) - f_{3t} (1 - p_t)^2 = 0; \tag{A 12j}$$

and for  $t = T$ :

$$\frac{\delta L}{\delta W_{B,T}} = -k_T + 2 = 0, \tag{A 12k}$$

$$\frac{\delta L}{\delta W_{b,T}} = -\lambda_T + 2 = 0, \tag{A 12l}$$

$$\frac{\delta L}{\delta p_T} = -\gamma_T + 2a = 0. \tag{A 12m}$$

Similar to before, equations (A 12a), (A 12b) and (A 12c), along with (A 12j), which results in constraint (12a), can be used to solve for optimum control variables in generation  $t$  given state variables for  $t$  and Lagrange multipliers for  $t + 1$ . Combining (A 12a) and (A 12b) gives:

$$\begin{aligned} x_{1t} &= \frac{\{ (k_{t+1} + \lambda_{t+1}) (\bar{A}_{B,t} + \bar{A}_{b,t} + x_{2t} \sigma) - 4k_{t+1} \bar{A}_{B,t} - 2\gamma_{t+1} \}}{2k_{t+1} \sigma}. \end{aligned} \tag{A 13a}$$

Combining (A 12b) and (A 12c) gives:

$$\begin{aligned} x_{3t} &= \frac{\{ (k_{t+1} - \lambda_{t+1}) (\bar{A}_{B,t} + \bar{A}_{b,t} + x_{2t} \sigma) - 4\lambda_{t+1} \bar{A}_{b,t} + 2\gamma_{t+1} \}}{2\lambda_{t+1} \sigma}. \end{aligned} \tag{A 13b}$$

Using (A 13a) and (A 13b), which express truncation points  $x_{1t}$  and  $x_{3t}$  in terms of  $x_{2t}$ , given state variables  $p_t$ ,  $W_{B,t}$  and  $W_{b,t}$ , and Lagrange multipliers for  $t + 1$ , and using (A 12j) (or (12a)), which constrains the overall fraction selected to  $Q$ , the iterative procedure of Ducrocq & Quaas (1988) can again be used to find optimum truncation points  $x_{mt}$  and proportions selected from each genotype class,  $f_{mt}$ .

Lagrange multipliers  $k_t$ ,  $\lambda_t$  and  $\gamma_t$ , for  $t = 0$  to  $T - 1$  are obtained from the following three sets of backward recursive equations:

For  $\gamma_t$ , from (A 12d), and substituting (A 12c) for  $\epsilon_t$ :

$$\begin{aligned} \gamma_t &= \frac{1}{2Q} \{ k_{t+1} \{ (2f_{1t} - f_{2t}) W_{B,t} + f_{2t} W_{b,t} \\ &+ [2p_t z_{1t} + (1 - 2p_t) z_{2t}] \sigma \} + \lambda_{t+1} \{ (-2f_{3t} + f_{2t}) W_{b,t} \\ &- f_{2t} W_{B,t} + [2(p_t - 1) z_{3t} + (1 - 2p_t) z_{2t}] \sigma \} \\ &+ \gamma_{t+1} \{ 4f_{1t} p_t + 2f_{2t} (1 - 2p_t) \} \\ &- \lambda_{t+1} (2\bar{A}_{b,t} + x_{3t} \sigma) \\ &\times \{ 2f_{1t} p_t + 2f_{2t} (1 - 2p_t) + 2f_{3t} (p_t - 1) \} \}, \end{aligned} \tag{A 14a}$$

with as starting value from (A 12m):  $\gamma_T = 2a$ .

For  $k_t$ , from (A 12e):

$$k_t = \frac{1}{2Q} \{ k_{t+1} [2f_{1t} p_t + f_{2t} (1 - p_t)] + \lambda_{t+1} [f_{2t} (1 - p_t)] \}, \tag{A 14b}$$

with as starting value (from (A 12k)):  $k_T = -2$ .

For  $\lambda_t$ , from (A 12f):

$$\lambda_t = \frac{1}{2Q} \{ \lambda_{t+1} [2f_{3t}(1-p_t) + f_{2t}p_t] + k_{t+1} [f_{2t}p_t] \} \quad (\text{A } 14c)$$

with as starting value (from (A 12l)):  $\lambda_T = 2$ .

State variables are obtained from the following forward recursive equations:  $p_t$  from (A 12i) (which results in (12b)), with as starting value  $p_0$ ;  $\bar{W}_{B,t}$  from (A 12g) (which results in (12c)), with as starting value  $\bar{W}_{B,0}$ ; and  $W_{b,t}$  from (A 12h) (which results in (12d)), with as starting value  $W_{b,0}$ . Therefore, optimum solutions need to satisfy equations (A 13a, b), (A 14a–c) and (A 12g–j).

Based on the above, the following iterative procedure to find the optimum can be derived:

1. Set  $k_t = \lambda_t = 2$  and  $\gamma_t = 2a$  for all  $t$ .
2. For  $t = 0$  and given  $p_0, k_0, \lambda_0, \gamma_0, \bar{A}_{B,0}$  and  $\bar{A}_{b,0}$ , derive  $f_{m,0}$  that satisfy the constraint given by (A 12j) (overall fraction selected), using the truncation selection procedure of Durocq & Quaas (1988), with relationships between truncation points given by (A 13a) and (A 13b). Compute the associated values for  $z_{m,0}$  and  $i_{m,0}$ .
3. Compute  $p_{t+1}, \bar{W}_{B,t+1}$  and  $W_{b,t+1}$  based on (A 12i), (A 12g) and (A 12h).
4. Repeat steps 2 and 3 for  $t = 1$  to  $T-1$ .
5. Using equations (A 14a–c), recursively compute new values for  $k_t, \lambda_t$  and  $\gamma_t$ , given the solutions for  $p_t$  and  $f_{m,t}$  obtained from step 4, and given  $\lambda_T = k_T = 2$  and  $\gamma_T = 2a$ . A multiplicative relaxation factor may be required here, reducing changes in values for  $k, \lambda$  and  $\gamma$  from one iteration to another, to allow convergence.
6. Repeat steps 2 through 5 until  $k, \lambda$  and  $\gamma$  converge to a stable solution.
7. Compute optimum index weights based on (4a) and (4b).

Note that the first iteration in this procedure corresponds to genotypic selection.

This research was motivated by discussions with Dr John Gibson from the University of Guelph. We thank Dr W. G. Hill and an anonymous reviewer for constructive comments on the manuscript. Part of this research was conducted while J.C.M.D. was on leave at the Department of Animal Breeding of the Wageningen Agricultural University, the Netherlands. Financial support from the Natural Sciences and Engineering Research Council of Canada, the Wageningen Agricultural International Exchange fund, and the EU Human Capital and Mobility Research Network CHRX-CT94-0508 is gratefully acknowledged.

## References

Bryson, A. E. & Ho, Y.-C. (1975). *Applied Optimal Control: Optimization, Estimation and Control*. New York: Wiley.

- Bulmer, M. G. (1980). *The Mathematical Theory of Quantitative Genetics*. Oxford: Clarendon Press.
- Dekkers, J. C. M., Birke, P. V. & Gibson, J. P. (1995). Optimum linear selection indexes for multiple generation objectives with non-linear profit functions. *Animal Science* **65**, 165–175.
- Durocq, V. & Quaas, R. L. (1988). Prediction of genetic response to truncation selection across generations. *Journal of Dairy Science* **71**, 2543–2553.
- Eikelenboom, G. & Minkema, D. (1974). Prediction of pale, soft, exudative muscle with a non-lethal test for the halothane-induced porcine malignant hyperthermia syndrome. *Tijdschrift voor Diergeneeskunde* **99**, 421–426.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Harlow, UK: Longman.
- Fuji, J., Otsu, K., Zorzato, F., Deleon, S., Khanna, V. K., Weiler, J., O'Brien, P. J. & MacLennan, D. H. (1991). Identification of a mutation in the porcine ryanodine receptor that is associated with malignant hypothermia. *Science* **253**, 448–451.
- Gibson, J. P. (1994). Short-term gain at the expense of long-term response with selection of identified loci. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, **21**, 201–204.
- Henderson, C. R. (1988). Theoretical basis and computational methods for a number of different animal models. *Journal of Dairy Science* **71** (Suppl. 2), 1.
- Kamien, M. I. & Schwartz, N. L. (1981). *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. Amsterdam: North-Holland.
- Kashi, Y., Hallerman, E. & Soller, M. (1990). Marker-assisted selection of candidate bulls for progeny testing programmes. *Animal Production* **51**, 63.
- Kennedy, B. W., Quinton, M. & Van Arendonk, J. A. M. (1992). Estimation of effects of single genes on quantitative traits. *Journal of Animal Science* **70**, 2000–2012.
- Lewis, F. L. (1986). *Optimal Control*. New York: Wiley.
- Luo, Z. W., Thompson, R. & Woolliams, J. A. (1997). A population genetics model of marker-assisted selection. *Genetics* **146**, 1173–1183.
- Meuwissen, T. H. E. & Van Arendonk, J. A. M. (1992). Potential improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. *Journal of Dairy Science* **75**, 1651–1659.
- Ruane, J. & Colleau, J. J. (1995). Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genetical Research* **66**, 71–83.
- Ruane, J. & Colleau, J. J. (1996). Marker-assisted selection for a sex-limited character in a nucleus breeding population. *Journal of Dairy Science* **79**, 1666–1678.
- Shuster, D. E., Kehrli, Jr, M. E., Ackermann, M. R. & Gilbert, R. O. (1992). Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. *Proceedings of the National Academy of Sciences of the USA*, **89**, 9225.
- Smith, C. & Simpson, S. P. (1986). The use of genetic polymorphisms in livestock improvement. *Journal of Animal Breeding and Genetics* **103**, 205–217.
- Van Arendonk, J. A. M., Bovenhuis, H., Van Der Beek, S. & Groen, A. F. (1994). Detection and exploitation of markers linked to quantitative traits in farm animals. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, **21**, 193–200.
- Van Der Beek, S. & Van Arendonk, J. A. M. (1994). Marker assisted selection in a poultry breeding program. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, **21**, 237–240.

Woolliams, J. A. (1990). Strategies to maximise selection progress in dairy cattle. *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, Edinburgh, **14**, 15–24.

Woolliams, J. A. & Pong-Wong, R. (1995). Short-versus long-term responses in breeding schemes. *Book of Abstracts, European Association of Animal Production*, Prague.