

# Genotype-assisted optimum contribution selection to maximize selection response over a specified time period

THEO H. E. MEUWISSEN<sup>1\*</sup> AND ANNA K. SONESSON<sup>2</sup>

<sup>1</sup>Centre for Integrative Genetics, Institute of Animal Science, Agricultural University of Norway, PO Box 5025, 143 Ås, Norway

<sup>2</sup>AKVAFORSK, PO Box 5010, 1432 Ås, Norway

(Received 4 August 2003 and in revised form 11 December 2003 and 7 April 2004)

## Summary

Genotype-assisted selection (GAS), i.e. selection for an identified quantitative trait locus (QTL) and polygenic background genes, has been shown to increase short-term genetic gain but may reduce long-term genetic gains. In order to avoid this reduction of long-term gain, multi-generation optimization of truncation selection schemes is needed. This paper presents a multi-generation optimization of optimum contribution (OC) selection with selection on an identified QTL. This genotype-assisted optimum contribution (GAOC) selection method assumes that the optimum selection differential at the QTL is constant over the time horizon, and achieves this by controlling the increase of the frequency of the positive QTL allele. Implementation was straightforward by an additional linear restriction in the OC algorithm. GAOC achieved 35.2%, 2.3% and 1.1%, respectively, more cumulative genetic gain than OC selection (ignoring the QTL) using time horizons of 5, 10 and 15 generations. When one-generation optimization of GAS was used instead of multi-generation optimization, these figures were 2.8%, 3.1% and 3.2%, respectively. Simulated annealing was used to optimize the increases of the frequency of the positive QTL allele in order to test the optimality of GAOC. This latter resulted in genetic gains that were always within 0.4% of those of GAOC. In practice, short-term genetic gains are also important, which makes one-generation optimization of genetic gain closer to optimal.

## 1. Introduction

Genotype-assisted selection (GAS) is possible when (part of) the genetic variation is due to a detected gene, such as is the case for the K232A substitution in the *DGAT1* gene (Grisart *et al.*, 2002). In the future, more of the causal mutations underlying quantitative trait loci (QTLs) will be detected, and GAS will become an increasingly common tool for the genetic improvement of livestock. It is, however, not expected that all the mutations underlying genetic variation in quantitative traits will be detected, so that selection for 'background' genes (often termed polygenes) remains important.

GAS and, similarly, marker-assisted selection (MAS) have been found to increase short-term selection response, but, in the longer term (5–10

generations), the highest selection response was achieved by conventional selection, i.e. by ignoring the QTL (Gibson, 1994; Pong-Wong & Woolliams, 1998; Villanueva *et al.*, 1999, 2002*a, b*). This paradox, that ignorance about the existence of the QTL gives the highest selection response, has become known as the Gibson effect. In the long term, both GAS and conventional selection will fix the positive QTL allele, and thus both achieve maximum gain at the QTL. GAS will fix the positive QTL allele faster, and thus will achieve more early (short-term) selection response. The high emphasis on the QTL with GAS, however, results in a reduced early response at polygenes, which is not fully recovered during the later generations of selection due to the non-linear relationship between selection pressure and selection response (Pong-Wong & Woolliams, 1998).

Dekkers & Van Arendonk (1998) and Manfredi *et al.* (1998) solved the Gibson effect by optimizing the

\* Corresponding author. Tel: +47 64947974. Fax: +47 64947960.  
e-mail: theo.meuwissen@iha.nlh.no

weights of the QTL and polygenic estimated breeding values in the selection criterion over the planning horizon. Manfredi *et al.* (1998) also optimized matings between animals. In one-generation-GAS schemes, i.e. GAS schemes where equal weights are given to estimates of polygenic effects and of QTL genotype effects, the genetic gain in the next generation is maximized (the time horizon is one generation), but genetic gain is suboptimal if this one-generation scheme is repeatedly used over a period of e.g. 10 generations. Using optimal control theory and dynamic programming, the weights for the QTL genotype effects were optimized for every generation over a period of e.g. 10 generations. The latter requires the prediction of the selection response over the time horizon given any set of weights. This optimization of the weights of the QTL breeding value is possible for phenotypic and BLUP selection. Chakraborty *et al.* (2002) extended this approach to multiple QTLs.

Optimum contribution (OC) selection maximizes genetic gain while constraining the rate of inbreeding in the population by optimizing the contributions of the parents to the next generation (Meuwissen, 1997; Grundy *et al.*, 1998; Meuwissen & Sonesson, 1998). These authors also showed that OC selection achieves substantially more genetic gain at the same rate of inbreeding than truncation selection for BLUP breeding value estimates when breeding values are purely polygenic. When OC selection that ignored the QTL was applied to populations with segregating QTLs, it yielded more genetic gain than truncation selection schemes that account for the QTL in the short and the long term (Villanueva *et al.*, 1999, 2002a). But these authors also showed that the Gibson effect occurred for OC selection when used as a one-generation scheme (i.e. estimates of polygenic and QTL effect have equal weights). Villanueva *et al.* (2002b) therefore tried to combine optimal weighting of the QTL across generations, using the Dekkers & Van Arendonk (1998) approach, which assumes truncation selection, and OC selection. The main limitation of this approach is that the weights are derived under the assumption of truncation selection while they are used in an OC selection scheme, and are thus suboptimal. The derivation of optimal weights requires the deterministic prediction of genetic gains of OC selection, which is not yet possible.

The aim here is to develop a genotype-assisted optimum contribution (GAOC) selection method that maximizes the genetic gain over a specified period of time while controlling the inbreeding, and that resolves the Gibson effect (i.e. knowledge about the QTL improves genetic gain over the specified time period). The problem of determining optimal weights for the QTL across generations is avoided. Instead we control the increase in the frequency of the positive QTL

allele using OC selection. The dynamic nature of OC selection, i.e. it adapts to the selection candidates at hand, will be maintained in GAOC selection. GAOC and OC (ignoring the QTL) will be compared for their genetic gains over different time horizons by Monte Carlo simulations.

## 2. Methods

### (i) GAOC selection

OC selection is designed to maximize genetic gain in generation  $t$  with a constraint on the rate of inbreeding, i.e. the genetic level of generation  $t+1$  is (Meuwissen, 1997):

$$G_{t+1} = \mathbf{c}_t' \mathbf{EBV}_t, \quad (1)$$

where  $\mathbf{EBV}_t$  is a vector of genotype-assisted estimates of breeding values of the selection candidates in generation  $t$  (with equal weights for the polygenic breeding value estimate and QTL effect), and  $\mathbf{c}_t$  is a vector of optimal contributions of the selection candidates to the next generation, which is proportional to the fraction of the offspring that each candidate should have. In  $\mathbf{EBV}_t$ , the weights for the polygenic breeding value estimate and QTL effects are equal and not optimized as was done by Dekkers & Van Arendonk (1998) and Manfredi *et al.* (1998). In GAOC, the genetic response at the QTL will be limited directly by constraining the response at the QTL, as shown in the next section, instead of by reducing the weight of the QTL genotype in the total EBV. This approach circumvents the optimization of these weights.

The contributions are optimized subject to a quadratic constraint (2) and a set of linear constraints (3):

$$K_{t+1} = \mathbf{c}_t' \mathbf{A}_t \mathbf{c}_t \quad (2)$$

$$\mathbf{s} = \mathbf{Q}_t \mathbf{c}_t \quad (3)$$

where constraint (2) ensures that the average relationships of the selected parents,  $\mathbf{c}_t' \mathbf{A}_t \mathbf{c}_t$ , and thus the inbreeding coefficients of their offspring, follow a predefined path, which is defined by the value of the restriction at each point in time,  $K_t$ . In OC selection, the  $K_t$  values are chosen such that the rate of inbreeding,  $\Delta F$ , is constant (Grundy *et al.*, 1998), i.e.

$$K_{t+1} = K_t + (2 - K_t)\Delta F.$$

The linear constraints (3) were originally used to ensure that the contributions of the males and of the females sum to  $\frac{1}{2}$ , i.e.  $\mathbf{s} = [\frac{1}{2} \frac{1}{2}]'$ , and  $\mathbf{Q}_t$  is a  $(2 \times n)$  design matrix indicating the sex of each of  $n$  candidates (Meuwissen, 1997). However, additional linear constraints are easily added and will not change the general form of the optimal solution to the maximization of (1) under the constraints (2) and (3) (Grundy *et al.*, 1998). Here, we are particularly

interested in constraining the selection response at the QTL, since this is initially too high and results in too low a response at the polygenes (Gibson, 1994). The allele frequency of the positive QTL allele in generation  $t+1$  equals the frequency of the QTL in the parents of generation,  $t+1$ , weighted by the contribution of the parents, i.e.

$$q_{t+1} = \mathbf{q}_t' \mathbf{c}_t,$$

where  $\mathbf{q}_t$  is a vector of QTL allele frequencies of every candidate, i.e. elements of  $\mathbf{q}_t$  are 0,  $\frac{1}{2}$  or 1 if the candidate carries 0, 1, or 2 of the positive QTL alleles, respectively. Thus, if we do not want too much selection pressure at the QTL and want, for example,  $q_{t+1} \leq x$ , we can use the following GAOC algorithm to implement the constraint that  $\mathbf{q}_t' \mathbf{c}_t \leq x$ :

*Step 1:* Optimize the contributions,  $\mathbf{c}_t$ , while ignoring the constraint  $\mathbf{q}_t' \mathbf{c}_t \leq x$ , e.g. using the algorithm of Meuwissen (1997).

*Step 2:* If  $\mathbf{q}_t' \mathbf{c}_t \leq x$  and  $\mathbf{q}_t' \mathbf{c}_t > 0$ , accept the solution of Step 1 and finish. Otherwise, include the constraint  $\mathbf{q}_t' \mathbf{c}_t = x$  in the set of constraints (2), i.e. set:  $\mathbf{s} = [\frac{1}{2} \frac{1}{2} x]'$  and add an extra row to  $\mathbf{Q}_t$  containing the elements of  $\mathbf{q}_t'$  ( $\mathbf{q}_t' \mathbf{c}_t > 0$  ensures that the positive QTL allele is not lost.)

*Step 3:* Optimize the contributions,  $\mathbf{c}_t$ , with the additional constraint  $\mathbf{q}_t' \mathbf{c}_t = x$  included. Again the algorithm of Meuwissen (1997) can be used.

Step 1 is needed to implement the constraint  $\mathbf{q}_t' \mathbf{c}_t \leq x$ , while ignoring Step 1 would have implemented the constraint  $\mathbf{q}_t' \mathbf{c}_t = x$ , which might put more emphasis on the QTL than the one-generation OC scheme of Step 1 (which might not reach a QTL frequency as high as  $x$ ). In the following section, we determine the optimum path of the QTL allele, i.e. the optimum frequency that the QTL allele should have in each generation  $t$ .

### (ii) The optimum path of the QTL allele

When optimizing genetic gain for a specified time horizon and when ignoring linkage disequilibrium (LD) between the QTL and the polygenes, Dekkers & Van Arendonk (1998) noticed that, although the weights given to the QTL genotypes vary widely over generations, the intensity of selection at the QTL (and at the polygenes) is constant. For simplicity, we will ignore any linkage disequilibrium between the polygenes and the QTL, and develop an OC selection method that maintains constant selection intensity at the QTL, i.e.

$$\Delta q_t / \sigma_{q(t)} = C \quad (4)$$

where  $\Delta q_t$  is the selection differential at the QTL in generation  $t$ , which equals the change in allele

frequency;  $\sigma_{q(t)}$  is the standard deviation of the allele frequencies of the individuals; and  $C$  is the constant selection intensity at the QTL. Note that  $\sigma_{q(t)}^2$  is the variance due to an additive QTL in generation  $t$  with an effect of  $a = \frac{1}{2}$ , where  $a$  is the additive gene effect as defined by Falconer & Mackay (1996), i.e.  $\sigma_{q(t)}^2 = \frac{1}{2} q_t (1 - q_t)$ .

The optimal solution of Dekkers & Van Arendonk (1998) demonstrated that the positive QTL allele should be close to fixation at the end of the time horizon. This is probably because the heritability of the QTL polymorphism is 1, and this high heritability should be exploited as much as possible by depleting the variance at the QTL, but not too fast, since this would lead to reduced selection differentials for the polygenes during the generations where there is strong selection for the QTL. There will be increased selection differentials for the polygenes after the QTL allele has been fixed, but, because of the non-linear relationship between the selected fraction and selection intensity, this does not fully compensate for the loss in polygenic response during the earlier generations. Thus, the variance at the QTL should reach 0 at the end of the time horizon, such that the polygenic selection differential is constant. The above argument assumes that there is no relationship between the polygenes and the QTL genotype (no LD), otherwise it matters whether the polygenic response is achieved in the best QTL genotype class or in the worst. However, also when accounting for LD, Dekkers & Van Arendonk (1998) found that the optimum solution brought the QTL close to fixation at the end of the time horizon.

When the initial QTL allele frequency,  $q_0$ , and  $C$  from equation (4) are known, the following recursive equations define the increase in allele frequency per generation ( $\Delta q_t$ ):

$$\Delta q_t = C * \sigma_{q(t)} = C * \sqrt{[\frac{1}{2} q_t (1 - q_t)]} \quad (5)$$

$$q_{t+1} = q_t + \Delta q_t.$$

The initial QTL allele frequency,  $q_0$ , is assumed known here. By trial and error,  $C$  is found such that  $q_T = 0.99$ , i.e. the QTL is close to fixation at the end of the time horizon,  $T$ . The above  $q_t$  values define an optimal path for the QTL allele to fixation (optimal in the sense that it achieves a constant high intensity of selection at the QTL and brings the QTL allele close to fixation). These  $q_t$  values are used in Step 3 of the GAOC algorithm (see Section 2(ii)).

### (iii) The ANNEAL scheme

In order to test whether the path of the QTL allele used by GAOC is close to optimal, simulated annealing (e.g. Press *et al.*, 1989) was also used to optimize the increases in frequencies of the QTL allele in each generation while maximizing genetic response over

Table 1. Parameters of the simulated breeding schemes

No. of animals per generation (50% males; 50% females)	200
Constraint on rate of inbreeding	1% per generation
Bi-allelic QTL	
Initial frequency of positive allele	0.15
Additive effect (Falconer & Mackays (1996) definition)	1
Dominance effect	0
Polygenic variance	0.25
Environmental variance	5
Trait recording	Before selection
Time horizon ( $T$ )	5, 10 or 15 generations
Objective of breeding scheme	Maximum genetic level in generation $T$
No. of replicated simulations	
For scheme comparisons	50
Within the simulated annealing optimization	10 <sup>a</sup>

<sup>a</sup> In order to save computer time, the simulated annealing optimization algorithm used only 10 replicated simulations to estimate genetic gain of alternative breeding schemes. The optimal scheme that was found by simulated annealing was, however, simulated 50 times in order to obtain results as accurate as those for the other schemes.

the time horizon and constraining inbreeding. This scheme is denoted ANNEAL.

In the ANNEAL scheme, the simulated annealing algorithm was used to optimize the path of the QTL, i.e. the frequencies  $q_1, q_2, \dots, q_T$ , such that  $G_T$  is maximized, where  $G_T$  is evaluated by 10 replicated Monte Carlo simulations of the scheme. Given a (trial) set of frequencies  $q_1, q_2, \dots, q_T$ , the ANNEAL scheme used the GAOC algorithm to (1) achieve the indicated frequencies  $q_1, q_2, \dots, q_T$ ; (2) to optimize the contributions,  $c_i$ ; and (3) to constrain the rate of inbreeding.

Simulated annealing is a computationally intensive Monte-Carlo-based optimization algorithm, which implies here that for many 'random/trial sets of  $q_i$  values' the genetic gain,  $G_T$ , needed to be calculated. In the absence of a deterministic method, Monte Carlo simulation of the suggested breeding scheme was used to evaluate  $G_T$ . Because these Monte Carlo simulations needed to be conducted for very many sets of  $q_i$  values, the number of replicated simulations for each was limited to 10. Note that the simulated annealing optimization of the path of the QTL allele of these ANNEAL schemes accounts for the LD between the QTL and polygenes, because the Monte Carlo simulations of breeding schemes automatically account for this LD.

Due to their very high computational demands, the ANNEAL schemes are not considered a viable option in practical applications. Ideally, they should be re-optimized every generation, since the realized selection response may differ from the planned response due to sampling, but this was not possible with the available computer capacity.

#### (iv) The GAS+OC scheme

One further selection strategy was tested, denoted GAS+OC, in which genotype-assisted EBV were

estimated, assuming equal weights for the QTL and polygenic breeding value, and these EBV were used for OC selection (Villanueva *et al.*, 1999, 2002a; called 'GAS with optimized selection' by these authors). Hence, GAS+OC is a one-generation optimization algorithm, equivalent to using only Step 1 of the GAOC algorithm.

#### (v) Simulated breeding schemes

The breeding schemes described in Table 1 were simulated in order to compare them for total genetic gain and its distribution over the QTL and polygenes. There were 100 male and 100 female selection candidates per generation. The inbreeding at unlinked neutral loci was constrained to 1% per generation by using OC selection. The initial heritability due to the polygenes and the bi-allelic QTL was quite low, i.e. 0.0917, as it is expected that extra genetic gains due to MAS/GAS are highest for lowly heritable traits (e.g. Meuwissen & Goddard, 1996). (These bi-allelic QTL simulations may also approximate the situation of a multiallelic QTL, where the best alleles are simulated by the positive allele, and the average effect of the remaining alleles is simulated by the negative allele.) In the first generation, approximately half the genetic variance is due to the QTL and half is due to the polygenes. The genetic variance due to the QTL was chosen to be rather large, which increases the effect of knowledge versus ignorance about the QTL allele, and thus the Gibson effect will be exaggerated relative to situations with smaller effects. Polygenes and QTL alleles were in linkage equilibrium in the first generation, where polygenic effects for animal  $i$ ,  $g_{i(0)}$ , were sampled from  $N(0,0.25)$  and the number of positive QTL alleles was sampled from  $\text{Bin}(n=2; p=0.15)$ . The effects of the QTL alleles were additive, and each positive QTL allele increased the phenotype

by  $a=1$ . In later generations,  $g_{i(t)}$  was sampled from  $N[\frac{1}{2}(g_{s(t-1)} + g_{d(t-1)}); \frac{1}{2} * 0.25 * (1 - \frac{1}{2}F_s - \frac{1}{2}F_d)]$ , where  $F_s$  ( $F_d$ ) is the inbreeding coefficient of the sire (dam). A paternally (maternally) inherited QTL allele was sampled at random from the QTL alleles of the sire (dam). The effects of the QTL were assumed additive, and phenotypic records were obtained as  $p_{i(t)} = g_{i(t)} + 1 * (\text{number of positive QTL alleles}) + e_{i(t)}$ , where  $e_{i(t)}$  represents the environmental effect and was sampled from  $N(0,5)$ .

For the GAS schemes, BLUP-EBVs were calculated using QTL genotype as a fixed effect, and EBV<sub>*i*</sub> was the sum of the QTL genotype and the individual animal effect. For the non-GAS schemes, BLUP-EBV were calculated ignoring the information of the QTL with the genetic variance including that due to the QTL (assuming the initial frequency of 0.15). These BLUP-EBV were used in OC selection, which implies using only Step 1 of the GAOC algorithm.

### 3. Results

Figure 1 shows total genetic gain, genetic gain at the QTL and the polygenes, for schemes that maximized the genetic level in generation  $T=5$ , i.e. quite short term breeding schemes (although 5 generations may equal 20 years for some species). The positive QTL allele was not lost in any of the replicated Monte Carlo simulations underlying Figs 1, 2 and 3. In Fig. 1, a QTL allele frequency of only 0.64 was reached with non-GAS OC selection, and thus maximum response at the QTL was not achieved. The GAS schemes, i.e. GAOC, ANNEAL and GAS+OC, thus had the opportunity to achieve substantially more selection response at the QTL: they reached QTL allele frequencies of 0.99, 0.98 and 1.0, respectively, and yielded substantially more total selection response, 35.2%, 35.6% and 31.6%, respectively, than the non-GAS scheme. GAS+OC fixed the positive QTL allele in generation 3, and selected thereafter only for the polygenes. Although GAS+OC's superiority over non-GAS became smaller during generations 4 and 5, non-GAS did not overtake GAS+OC, and thus the Gibson effect did not occur over this short time period. The multi-generation optimization scheme GAOC did, however, achieve 2.8% more genetic gain than GAS+OC.

ANNEAL and GAOC gave very similar total selection responses (ANNEAL had 0.3% more response than GAOC) and quite similar paths to these, with some irregular differences that may well be due to sampling (Fig. 1). Sampling errors on the estimates of  $G_T$  make it hard for ANNEAL to find the truly optimal path, especially when  $G_T$  is robust against deviations from the optimal path, for the differences in path between ANNEAL and GAOC lead to a very small difference in  $G_5$ .

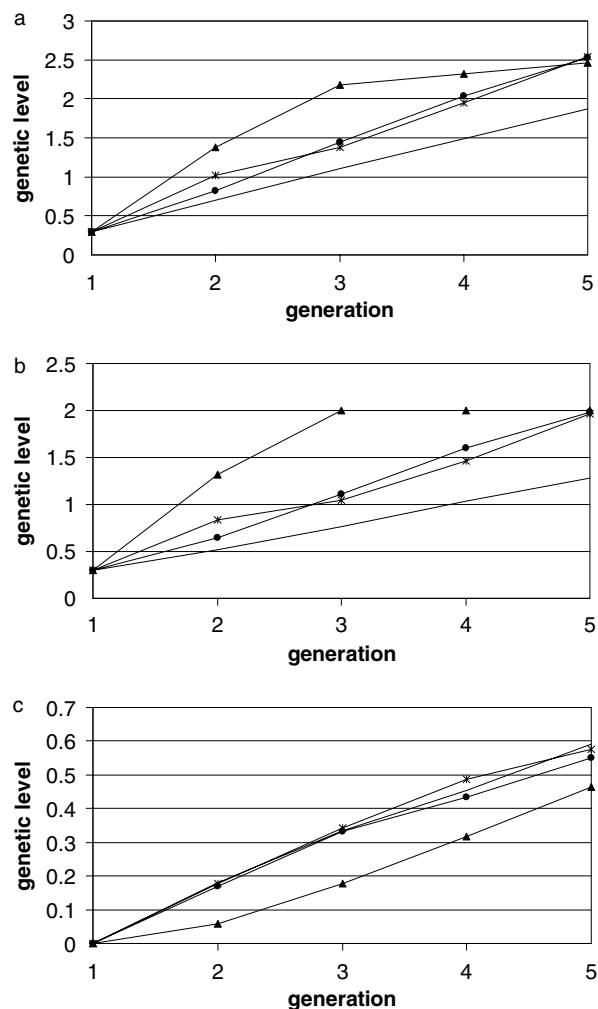


Fig. 1. Total (a), QTL (b) and polygenic (c) genetic gain when the time horizon is 5 generations for non-GAS (—), GAOC (●), GAS+OC (▲) and ANNEAL (×) schemes (averaged over 50 replicated Monte Carlo simulations).

When the time horizon was 10 generations (Fig. 2), all breeding schemes either fixed the positive QTL allele or came close to fixation of this allele (the allele frequency was the lowest for non-GAS with a frequency of 0.978). GAS+OC achieved the lowest total genetic gain, followed by non-GAS, GAOC and ANNEAL, which yielded 0.9%, 3.1% and 3.6% more, respectively. Here, the Gibson effect was evident and non-GAS yielded more total genetic gain than GAS+OC, i.e. one-generation optimization.

ANNEAL achieved 0.4% more genetic gain than GAOC, but the paths towards these selection responses were quite different: ANNEAL achieved substantially more selection response at the QTL during early generations and maintained some of that advantage until the end of the time horizon. The difference in paths is larger than in Fig. 1, and may be because more parameters are simultaneously

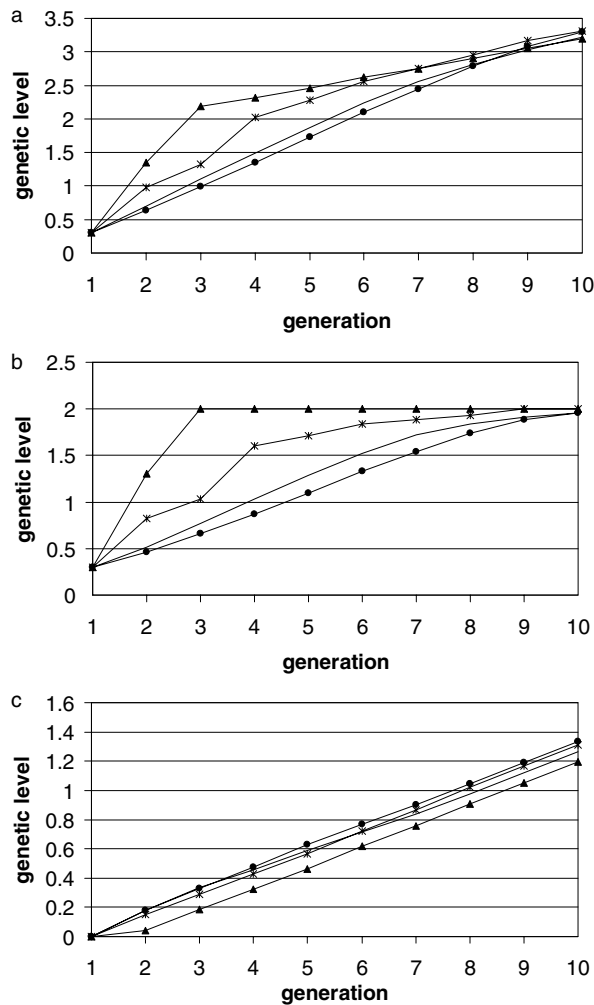


Fig. 2. Total (a), QTL (b) and polygenic (c) genetic gain when the time horizon is 10 generations for non-GAS (—), GAOC (●), GAS+OC (▲) and ANNEAL (×) schemes (averaged over 50 replicated Monte Carlo simulations).

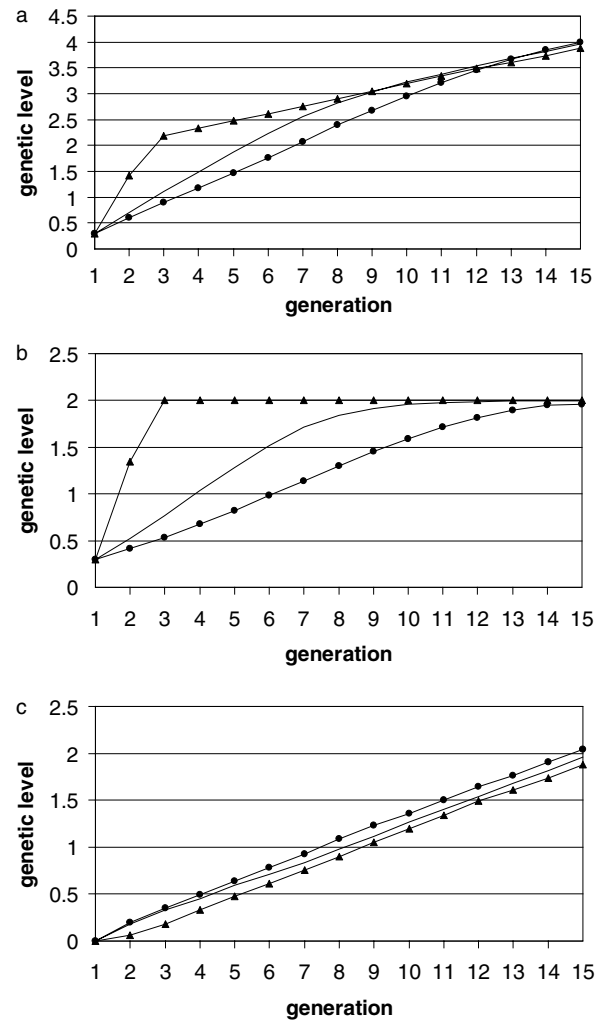


Fig. 3. Total (a), QTL (b) and polygenic (c) genetic gain when the time horizon is 15 generations for non-GAS (—), GAOC (●), GAS+OC (▲) and ANNEAL (×) schemes (averaged over 50 replicated Monte Carlo simulations).

optimized and the more long term objective leaves room for larger deviations between paths with similar  $G_{10}$ .

When the time horizon was 15 generations (Fig. 3), we could not perform the simulated annealing optimization due to the large number of parameters to optimize (15  $\Delta q_t$  values) and the long computer time needed per simulated scheme. Only the GAS+OC and non-GAS schemes achieved fixation of the QTL around generations 3 and 10, respectively, and selected thereafter for the polygenes, whereas GAOC came close to fixation only at generation 15. The selection pressure on the polygenes is therefore more uneven in the GAS+OC and non-GAS schemes (relatively low at the beginning when the QTL is still segregating and relatively high after generations 3 and 10, respectively, when the QTL is fixed) than in the GAOC scheme, which gave a more even selection pressure at the QTL and a more linear selection response for both the QTL and the total response. The result is that

GAS+OC yielded the lowest response, followed by non-GAS and GAOC, which yielded 2.1% and 3.2% more, respectively. These extra responses are small, but show that if the time horizon is sufficiently long, although the non-GAS schemes ignore the QTL they also can implicitly put too much emphasis on it, and can be improved by having more even selection pressure at the QTL and at the polygenes over the entire selection period.

#### 4. Discussion

In the long term, both GAS and conventional selection achieve maximum genetic gain at the QTL, provided the positive QTL is not lost (which was not the case for the large QTL considered here). Thus differences in total genetic gain are due to differences in polygenic response, which are usually small, i.e. differences of up to 5% more long-term response for

conventional selection were found in the literature (e.g. Pong-Wong & Woolliams, 1998). It is, however, rather unsatisfactory that conventional selection yields the highest genetic gain in the (medium to) long term, because most MAS and GAS schemes only optimize the selection response in the next generation – hence, the development of multi-generation optimization schemes such as those of Dekkers & Van Arendonk (1998), Manfredi *et al.* (1998), Villanueva *et al.* (2002*b*) and GAOC. Similar to GAOC, the method of Villanueva *et al.* (2002*b*) also combines OC selection with multi-generation optimization of GAS, but the approach was different, using Dekkers & Van Arendonk's method to obtain weights for the QTL breeding value assuming truncation selection while using these weights in OC selection. As these weights from truncation selection will be suboptimal when OC selection is applied, the method could be improved if they could be derived assuming OC selection. The latter requires deterministic prediction of genetic gain under OC selection, which is not yet possible, and maximization of these genetic gains.

In the GAOC approach, prediction of genetic gain was avoided by assuming that the selection differential at the QTL is constant over generations, as was found by Dekkers & Van Arendonk (1998) when ignoring linkage disequilibria between the QTL and the polygenes. Since these linkage disequilibria will occur, GAOC must be seen as an approximation to the true simultaneous optimization of OC selection and multi-generation GAS. The ANNEAL algorithm, however, does not assume linkage equilibrium, but gives very similar genetic gains to GAOC at the end of the time horizon. The medium-term gains were, though, often different from those of GAOC, especially for 10 generations (Fig. 2). Thus it may be concluded that the optimum for the  $\Delta q_t$  values is rather flat (quite different  $\Delta q_t$  can lead to very similar genetic gain), and that GAOC leads to close to optimal selection responses (simulated annealing did not give significant improvement; Figs 1, 2). An advantage of GAOC is that it is rather easy to implement: it requires determining the optimum path of the positive QTL allele using equation (5), and the implementation of the OC algorithm is very similar to conventional OC selection.

When compared at generation 10, GAOC achieved 3.1% more genetic gain than GAS + OC, and a similar comparison in the schemes of Villanueva *et al.* (2002*b*) resulted in 2.3% more genetic gain for their GAS<sub>OPT</sub> scheme than GAS<sub>STA</sub>. The main difference between these two comparisons is that Villanueva *et al.* maximized  $\Sigma = G_1 + \dots + G_{10}$  instead of  $G_{10}$ , where  $\Sigma$  also puts emphasis on medium-term responses. The maximization of  $G_{10}$  resulted in GAOC schemes that were superior in generation 10, but not earlier (Fig. 2), whilst the use of  $\Sigma$  prevents this. The

ANNEAL scheme of Fig. 2 was more superior in earlier generations. This may be due to the following possibilities: (i) a combination of chance (ANNEAL may not have found the optimal path for the QTL frequencies) and robustness of  $G_{10}$  to this path (which makes the truly optimal path hard to find); and/or (ii) accounting for LD between QTL and polygenes may result in more early generation selection response. Focus was here on  $G_T$  because it shows the Gibson effect most clearly, whereas earlier-generation responses are less likely to show it. For example, Fig. 1 did not show the Gibson effect (non-GAS yielded less  $G_5$  than GAS + OC), and  $\Sigma$  did not show it in Villanueva *et al.*'s results (PHE yielded less  $\Sigma$  than GAS<sub>STA</sub>).

Although GAOC uses EBV that give equal weight to QTL and polygenic breeding values, we believe it is a truly multi-generation optimization scheme, because:

1. GAOC looks ahead, in that it does not fix the positive QTL allele long before the time horizon is reached, using two results from Dekkers & Van Arendonk (1998), namely that the optimum selection intensity at the QTL is approximately constant, and that the QTL reaches near-fixation at the end of the time horizon. However, this ignores the fact that the intensity of selection should increase as the scheme approaches the end of the time horizon, when the scheme can maximize one-generation genetic gain.
2. GAOC controls the increase in frequency of the positive QTL allele, within which it optimizes the fraction selected from each QTL genotype. The latter is equivalent to optimizing the weight given to the QTL and polygenic breeding values (Dekkers & Van Arendonk, 1998).
3. In view of point 2, it may also be noted that the relative weights given to polygenic and QTL breeding values in the  $\mathbf{EBV}_t$  vector do *not* affect the optimum contributions,  $c_t$ , in the GAOC algorithm. This is because the genetic gain at the QTL is constrained and thus can not be altered by GAOC (all feasible solutions give the same genetic gain at the QTL; GAOC attempts to find the feasible solution that gives maximum gain for the polygenes). This point only holds if Step 3 of the GAOC algorithm is effective; but if the algorithm stops after Step 1, one-generation maximization of genetic gain does not yield over-large increases in the QTL allele frequency, and the weights for the QTL and polygenes should be equal. Hence, Step 1 of GAOC is important when one-generation optimization is not giving too early a fixation of the QTL, i.e. when the time horizon is short, and safeguards against putting too much selection pressure at the QTL. Points 1, 2 and 3 hold also for

the ANNEAL schemes, since they differ from GAOC only in the method for determining the optimum path of the QTL allele frequency.

In order to make the Gibson effect as large as possible here, the schemes were optimized and compared for maximum genetic level at the end of the time horizon. In practice, medium-term genetic gains are also of economic importance. Although suboptimal in all situations for the genetic gain at the end of the time horizon, GAS+OC gives substantially higher genetic levels before the time horizon is reached (Figs 1, 2, 3). The multi-generation optimization schemes can also be used to optimize an objective that values medium-term selection responses such as net present value (*NPV*) (Dekkers & Chakraborty, 2001), which values short-term gains higher than long-term gains:

$$NPV = \sum_{t=1}^T G_t * \frac{1}{(1+r)^t}$$

where  $G_t$  is the genetic level at generation  $t$ ; and  $r$  is the discount rate per generation, i.e. the depreciation of economic returns over time (Weller, 1994). *NPV* may be seen as the weighted sum of  $T$  time horizons, i.e.  $T$  maximizations of  $G_T$ . Hence, *NPV* in GAOC schemes can be maximized by determining  $\Delta q_t$  values for every time horizon, and using the weighted mean of these  $\Delta q_t$  values in the GAOC scheme, where the weights are  $1/(1+r)^t$ . This assumes an approximately linear relationship between  $\Delta q_t$  and  $G_t$  values, which may be justified because  $G_t$  was found to be rather robust against changes in  $\Delta q_t$  (Figs 1, 2). However, maximization of *NPV* will yield more short-term genetic gain, and thus the optimal schemes become closer to GAS+OC. Thus, its inferiority at the end of the time horizon of up to 3.6% should be seen as a maximum, and in practice the inferiority of GAS+OC is expected to be substantially less.

The genetic model used here was oversimplified in that there was one large QTL and all the other genes had small effects. The model is not realistic, especially, when considering long time horizons of 10 or 15 generations, which represent up to 50 years in cattle, and we hope to detect many more QTL in such a long period. This also argues for the use of GAS+OC, since it increases the chance that one QTL is near fixation before the next QTL is detected, and becomes the subject of GAS.

Helpful comments and suggestions from the reviewers and the Editor are gratefully acknowledged.

## References

- Chakraborty, R., Moreau, L. & Dekkers, J. C. M. (2002). A method to optimize selection on multiple identified quantitative trait loci. *Genetics, Selection, Evolution* **34**, 145–170.
- Dekkers, J. C. M. & Chakraborty, R. (2001). Potential gain from optimizing multigeneration selection on an identified quantitative trait locus. *Journal of Animal Science* **12**, 2975–2990.
- Dekkers, J. C. M. & van Arendonk, J. A. M. (1998). Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genetical Research* **71**, 257–275.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow, UK: Longman.
- Gibson, J. P. (1994). Short term gain at the expense of long term response with selection on identified loci. In *Proceedings of the Fifth World Congress on Genetics Applied to Livestock Production*, 7–12 August, Guelph, vol. 21, pp. 201–204. Guelph, Ontario, Canada: University of Guelph.
- Grisart, B., Coppetiers, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. & Snell, R. (2002). Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition. *Genome Research* **12**, 222–231.
- Grundy, B., Villanueva, B. & Woolliams, J. A. (1998). Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genetical Research* **72**, 159–168.
- Manfredi, E., Barbieri, M., Fournet, F. & Elsen, J. M. (1998). A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genetics, Selection, Evolution* **30**, 127–148.
- Meuwissen, T. H. E. (1997). Maximising the response of selection with a predefined rate of inbreeding. *Journal of Animal Science* **75**, 934–940.
- Meuwissen, T. H. E. & Goddard, M. E. (1996). The use of marker haplotypes in animal breeding schemes. *Genetics, Selection, Evolution* **28**, 161–176.
- Meuwissen, T. H. E. & Sonesson, A. K. (1998). Maximizing the response of selection with a predefined rate of inbreeding: overlapping generations. *Journal of Animal Science* **76**, 2575–2583.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1989). Minimization and maximization of functions. In *Numerical Recipes: The Art of Scientific Computing*, pp. 274–334. Cambridge: Cambridge University Press.
- Pong-Wong, R. & Woolliams, J. A. (1998). Response to mass selection when an identified major gene is segregating. *Genetics, Selection, Evolution* **30**, 313–337.
- Villanueva, B., Pong-Wong, R., Grundy, B. & Woolliams, J. A. (1999). Potential benefit from using an identified major gene in BLUP evaluation with truncation and optimal selection. *Genetics, Selection, Evolution* **31**, 115–133.
- Villanueva, B., Pong-Wong, R. & Woolliams, J. A. (2002a). Marker assisted selection with optimized contributions of the candidates to selection. *Genetics, Selection, Evolution* **34**, 679–703.
- Villanueva, B., Dekkers, J. C. M., Woolliams, J. A. & Settar, P. (2002b). Maximising genetic gain with QTL information and control of inbreeding. In *Proceedings of the Seventh World Congress on Genetics Applied to Livestock Production*, 19–23 August, Montpellier, France. Communication no. 22–18.
- Weller, J. I. (1994). *Economic Aspects of Animal Breeding*. London: Chapman and Hall.