

# Potential benefit from using an identified major gene in BLUP evaluation with truncation and optimal selection

Beatriz Villanueva<sup>a\*</sup>, Ricardo Pong-Wong<sup>b</sup>,  
Brian Grundy<sup>a</sup>, John A. Woolliams<sup>b</sup>

<sup>a</sup> Scottish Agricultural College, West Mains Road, Edinburgh,  
EH9 3JG, Scotland, UK

<sup>b</sup> Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, Scotland, UK

(Received 9 June 1998; accepted 28 January 1999)

**Abstract** – This study investigates the benefit of including information on an identified major gene in the estimation of breeding values in BLUP selection programmes. Selection for a quantitative trait is controlled by polygenes and a major locus with known effect. The benefit of using the gene information obtained in the short-term was maintained in the long-term by applying a selection tool which makes use of BLUP evaluation and optimisation of genetic contributions for maximising genetic gain while restricting the rate of inbreeding. In the mixed inheritance model the selection tool, initially proposed for an infinitesimal model, was able to restrict the rate of inbreeding to the desired value and to give higher rates of response than standard truncation selection both when using and ignoring the information on the major gene. The simple use of BLUP (standard truncation selection) allowed long-term benefits from using the gene in situations where the favourable allele was recessive or additive with large effect. © Inra/Elsevier, Paris

**major gene / optimal selection / BLUP selection / restricted inbreeding**

**Résumé** – Bénéfice possible de l'utilisation d'un gène majeur identifié dans une évaluation BLUP lors d'une sélection par troncature ou optimisée. Cette étude analyse le bénéfice pour la sélection, d'inclure l'information relative à un gène majeur identifié, dans l'estimation des valeurs génétiques par BLUP. La sélection porte sur un caractère quantitatif contrôlé par des polygènes et un locus majeur

\* Correspondence and reprints: Genetics & Reproduction Department, Animal Biology Division, SAC, Bush Estate, Penicuik, Midlothian EH26 0PH, Scotland, UK  
E-mail: b.villanueva@ed.sac.ac.uk

à effet connu. Le bénéfice à court terme de l'utilisation de l'information génique est maintenu à long terme grâce à un outil de sélection qui utilise le BLUP et qui optimise les contributions génétiques en vue de l'accroissement du progrès génétique à taux constant de consanguinité. Dans le modèle d'hérédité mixte, l'outil de sélection initialement proposé pour un modèle infinitésimal a été capable de restreindre le taux de consanguinité à la valeur désirée et de donner des taux de réponse plus élevés que la sélection classique par troncature, que l'on utilise ou que l'on ignore l'information sur le gène majeur. L'utilisation classique du BLUP (sélection standard par troncature) ne permet des bénéfices à long terme que si l'allèle favorable est récessif ou additif avec un effet important. © Inra/Elsevier, Paris

**gène majeur / sélection optimisée / sélection BLUP / taux de consanguinité contraint**

## 1. INTRODUCTION

Increasing numbers of single genes with large effect controlling quantitative traits are being identified in livestock species (e.g. Booroola and Callipyge genes in sheep, 'halothane' gene in pigs and 'double-muscling' gene in cattle) and this is expected to continue in the future. Genotyping of animals for particular genes is expensive but it may be cost effective if this information is used in selection programmes to produce additional and more targeted genetic response.

Studies evaluating the use of a major gene in mixed inheritance models for increasing genetic gain in mass selection programmes suggest a conflict between short- and long-term gains [5, 6, 11, 12, 15, 18]. Although in these studies the use of the available information on the major gene led to greater total genetic gain during the initial generations of selection, the accumulated response was lower when using genotype information by the time the favourable allele was fixed in both schemes (the scheme using the gene and the scheme ignoring the gene). The detrimental long-term effect was only avoided when the favourable allele was recessive with a large effect [12, 15]. The lower accumulated response when using genotype information in mass selection was mainly due to a decrease in the selection pressure applied to the polygenic background and, to a lesser extent, to a higher rate of inbreeding.

Many current breeding programmes use advanced technologies for estimating polygenic breeding values of the candidates for selection. When an infinitesimal model is assumed, animals are often selected on their BLUP (best linear unbiased prediction) estimated breeding values rather than simply on their phenotypic values. Standard selection based upon choosing individuals with the highest BLUP breeding values leads to increased inbreeding rates. This can be exacerbated if information on the major gene is used in the estimation of breeding values, as families associated with the most favourable genotype would contribute more to subsequent generations [15].

Recent developments in selection algorithms using BLUP breeding values allow the optimisation of selection decisions for giving maximum genetic gain over several generations of selection while restricting the rate of inbreeding to specific values [10, 14]. These procedures, initially proposed for the infinitesimal model, are very useful when comparing the efficiency of different schemes (at the same level of inbreeding) in the long-term.

The objective of this study was to investigate, using stochastic simulation, the value of including genotype information for an identified major gene in the estimation of breeding values for increasing short- and long-term genetic response in selection programmes using BLUP. Both standard truncation selection on BLUP breeding values and optimal selection for maximising gain while restricting inbreeding (which also uses BLUP) were considered. The efficiency of optimal selection for maximising gain and for restricting inbreeding in mixed inheritance models was also investigated.

## 2. METHODS

Monte Carlo simulations were used to compare schemes using or ignoring information on the major gene when the estimated breeding values (EBVs) are obtained from BLUP. Two selection procedures were considered: i) standard truncation selection with fixed number of parents and family sizes; and ii) 'optimal selection' in which the numbers of parents and their contributions are optimised each generation to maximise genetic gain while restricting the rate of inbreeding [10]. This optimisation differs from those described by Dekkers and van Arendonk [2] and by Manfredi et al. [13] where the purpose of the optimisation was to achieve the right emphasis given to the major gene relative to the polygenes for maximising gain without restrictions on inbreeding.

Comparisons of schemes were carried out in terms of short- and long-term accumulated genetic progress and inbreeding. A minimum of 200 replicates was run for each simulation.

### 2.1. Genetic model

The trait under selection was assumed to be genetically controlled by an infinite number of additive loci, each with infinitesimal effect (polygenes) plus a single biallelic locus (alleles A and B) with a major effect (major gene). The total genetic value of the  $i$ th individual was  $g_i = v_i + u_i$ , where  $v_i$  is the genotypic value due to the major locus and  $u_i$  is the polygenic effect. The major locus had an additive effect ( $a$ ), defined as half the difference between the two homozygotes, and a dominance effect ( $d$ ) defined as the difference between the heterozygote and the average of the two homozygotes. Thus, the genotypic value due to the major locus was  $a$ ,  $d$  and  $-a$  for individuals with genotype AA, AB and BB, respectively [3]. The additive variance explained by the major gene in the base population was  $\sigma_v^2 = 2p(1-p)\alpha^2$ , where  $p$  is the initial frequency of the favourable allele (A) and  $\alpha$  is the average effect of the gene substitution [3].

### 2.2. Simulation of the population

The base population ( $t = 0$ ) was composed of  $N = 120$  (60 males and 60 females) unrelated individuals. Generation 1 ( $t = 1$ ) was obtained from the mating of individuals selected at  $t = 0$ . The number of selection candidates ( $N$ ) was kept constant across 20 discrete generations of selection. The polygenic effect for animals of the base population was obtained from a normal distribution with mean zero and variance  $\sigma_v^2$ . The alleles at the major locus were chosen at random with  $p$  probability of an allele being the favourable A

(i.e. Hardy-Weinberg equilibrium is assumed). The phenotypic value for an individual  $i$  ( $y_i$ ) was obtained by adding to the total genetic value ( $g_i$ ) a normally distributed environmental component with mean zero and variance  $\sigma_e^2$ .

In subsequent generations, the polygenic effect of the offspring was generated as the average of the polygenic effects of their parents plus a random Mendelian deviation. The latter was sampled from a normal distribution with mean zero and variance  $(\sigma_u^2/2)[1 - (F_s + F_d)/2]$ , where  $F_s$  and  $F_d$  are the inbreeding coefficients of the sire and dam, respectively. The genotype for the major locus for each individual was obtained by sampling, at random, one allele from each parent.

### 2.3. Estimation of breeding values

In schemes using information on the major gene (genotype information) it was assumed that all individuals have known genotype for the major gene and that its effect was known without error.

When the information on the major locus was considered the selection criterion was

$$EBV_i = BLUP_i^* + w_i$$

where  $BLUP_i^*$  is the estimate of the polygenic breeding value for individual  $i$  and  $w_i$  is the breeding value due to the major locus effect. The estimate of the polygenic value was obtained from standard BLUP using the polygenic variance ( $\sigma_u^2$ ) and the total phenotypic values corrected for the major gene effect ( $y_i^* = y_i - v_i$ ). The breeding value of the single locus was  $2(1-p)\alpha$ ,  $[(1-p)-p]\alpha$  and  $-2p\alpha$  for individuals with genotype AA, AB and BB, respectively [3]. The frequency  $p$  and  $\alpha$  were updated each generation to obtain the breeding values.

When the information on the major locus was ignored the selection criterion was

$$EBV_i = BLUP_i$$

where  $BLUP_i$  is the estimated breeding value obtained from standard BLUP using the total genetic additive variance ( $\sigma_v^2 + \sigma_u^2$ ) of the base population and the phenotypic values ( $y_i$ ) uncorrected for the major gene effect.

Although the main objective of the study was to investigate the impact of using genotype information in BLUP-based selection methods, some schemes using mass selection were also simulated for comparison. When the genotype information was used in mass selection, the selection criterion was  $EBV_i = [(y_i - v_i)h_u^2] + w_i$ , where  $h_u^2$  is the polygenic heritability in the base population ( $h_u^2 = \sigma_u^2/(\sigma_u^2 + \sigma_e^2)$ ) [15]. When the information on the major locus was ignored, selection was carried out on uncorrected phenotypic values ( $EBV_i = y_i$ ).

### 2.4. Selection procedures

The benefit of including the information on the major gene in the estimation of breeding values was evaluated using either standard truncation selection or optimal selection [10]. The first case is static with fixed numbers of parents, whereas the second case is dynamic with a constraint on the rate of inbreeding.

### 2.4.1. Truncation selection

With standard truncation selection, a fixed number of individuals ( $N_s = 10$  males and  $N_d = 20$  females) with the highest estimated breeding values were selected to be parents of the next generation. Matings were hierarchical with each sire being mated at random to two dams and each dam producing three offspring of each sex.

### 2.4.2. Optimal selection

Optimal selection is a dynamic selection procedure in which the numbers of individuals selected and their contributions are not fixed but they are optimised for maximising genetic progress while restricting the rate of inbreeding to a specific value each generation. The procedure, initially proposed for an infinitesimal model, uses BLUP breeding values and the augmented numerator relationship matrix [10] to give the optimal selection decisions. A short description of the algorithm used for finding the optimal numbers of parents selected each generation and their optimal contributions is given in the Appendix. For a more detailed explanation of the method see Grundy et al. [10]. The EBVs used in the optimisation algorithm were those described in section 2.3 (i.e. not only the polygenic effects but also the major gene were considered in the optimisation). When optimal mass selection was simulated, the augmented numerator relationship matrix was still used to restrict the rate of inbreeding.

The solutions obtained with this algorithm are expressed as mating proportions (genetic contributions to the next generation) which sum to a half for each sex. The optimal number of offspring for individual  $i$  is  $2Nc_i$  (a real number), where  $c_i$  is the optimal solution (mating proportion) for individual  $i$ . The actual (integer) number of offspring for each parent was obtained as described in Grundy et al. [10]. Each parent was randomly allocated to different mates (among the selected individuals) to produce its offspring.

## 2.5. Parameters studied

The polygenic and the environmental variances were  $\sigma_u^2 = 0.2$  and  $\sigma_e^2 = 0.8$ , respectively, giving a polygenic heritability of 0.2. When the effect of the major gene was completely additive ( $d = 0$ ) different values for  $a$  were considered and results are presented for  $a = 0.5$ ,  $a = 1.0$  and  $a = 2.0$ . The initial frequency of the favourable allele was 0.15. Thus, at  $t = 0$ , the additive variance explained by the major locus and the total heritability were  $\sigma_v^2 = 0.06$ , 0.26 and 1.02 and  $h_t^2 = 0.25$ , 0.36 and 0.60 for  $a = 0.5$ , 1.0 and 2.0, respectively. These combinations of parameters avoided the loss of the favourable allele in all replicates, both in methods using and ignoring genotype information. Cases where the favourable allele was completely recessive ( $d = -a = -0.5$  and  $d = -a = -2.0$ ) were also studied. With recessive alleles some replicates lost the beneficial allele as described later.

### 3. RESULTS

#### 3.1. Truncation selection

Table I shows a comparison of genetic progress obtained when ignoring ( $I_T$ ) and using ( $G_T$ ) the information on the major gene in the selection criterion with truncation BLUP selection. The effect of the major gene was completely additive (i.e.  $d = 0$ ). The scheme using the individuals' genotypes yielded a greater total genetic gain than the scheme ignoring the genotype in the initial generations of selection, while the major locus was still segregating. However, at the time when the favourable allele is fixed in both  $I_T$  and  $G_T$ , the accumulated genetic response was lower with  $G_T$  when the major locus has a moderate effect ( $a = 0.5$ ). Fixation of the favourable allele occurred after six generations of selection in scheme  $G_T$  but after 18 generations with scheme  $I_T$  (although by generation 8 the frequency of the favourable allele was already higher than 0.95). At  $t = 18$  the gain obtained when using the genotype information was around 2 % lower than the gain obtained when ignoring that information.

**Table I.** Total accumulated genetic gain and frequency of the favourable allele ( $p$ ) over generations ( $t$ ) obtained from truncation BLUP selection ignoring ( $I_T$ ) or using ( $G_T$ ) genotype information on a major additive gene of moderate ( $a = 0.5$ ) or large effect ( $a = 2.0$ ). Genetic gains obtained from  $G_T$  are expressed as deviations from those obtained from  $I_T$ . The initial  $p$  was 0.15.<sup>a</sup>

$t$	$a = 0.5$				$a = 2.0$			
	Total accumulated gain		$p$		Total accumulated gain		$p$	
	$I_T$	$G_T - I_T$	$I_T$	$G_T$	$I_T$	$G_T - I_T$	$I_T$	$G_T$
1	0.339	0.115	0.24	0.47	1.309	0.190	0.44	0.50
2	0.739	0.225	0.38	0.82	2.879	0.404	0.80	0.93
3	1.117	0.232	0.53	0.98	3.778	0.072	0.99	1.00
4	1.491	0.143	0.68	1.00	4.075	0.053	1.00	1.00
5	1.828	0.064	0.81	1.00	4.313	0.066	1.00	1.00
10	3.097	-0.090	0.99	1.00	5.409	0.067	1.00	1.00
20	4.769	-0.100	1.00	1.00	7.252	-0.112	1.00	1.00

<sup>a</sup> Standard errors ranged from 0.0034 to 0.0281 for total genetic values and from 0.0 to 0.01 for frequency of the favourable allele.

On the other hand, when the major gene had a larger effect ( $a = 2.0$ ), the advantage of using the individuals' genotype observed in the early generations was maintained for several generations after the favourable allele was fixed in both schemes. Fixation of the favourable allele occurred after only three generations of selection and at  $t = 4$  the advantage of  $G_T$  over  $I_T$  was around 1 %.

The lower gain in the medium- and long-term obtained with  $G_T$  and  $a = 0.5$  was due to the faster increase in the frequency of the favourable allele which led to a lower polygenic gain in the early generations, when the major gene was still segregating (table II). The highest rate of response in the polygenic component

was obtained when the favourable allele was fixed. With  $a = 0.5$ , the initial loss in polygenic gain with  $G_T$  was not compensated for in later generations by the higher accuracy in estimating the EBVs with this method, and the result was that the use of the gene led to a decreased gain. With  $a = 2.0$  there was also a reduction in the rate of polygenic gain in the early generations before fixation but the increased accuracy when using the genotype information compensated for the initial loss in polygenic gain.

**Table II.** Rate of polygenic gain and inbreeding coefficient per generation ( $t$ ) obtained from truncation BLUP selection ignoring ( $I_T$ ) or using ( $G_T$ ) genotype information on a major additive gene of moderate ( $a = 0.5$ ) or large effect ( $a = 2.0$ ).<sup>a</sup>

$t$	$a = 0.5$				$a = 2.0$			
	Rate of polygenic gain <sup>b</sup>		Inbreeding		Rate of polygenic gain		Inbreeding	
	$I_T$	$G_T$	$I_T$	$G_T$	$I_T$	$G_T$	$I_T$	$G_T$
1	0.248	0.135	0.000	0.000	0.163	0.092	0.000	0.000
2	0.264	0.162	0.041	0.030	0.097	0.077	0.034	0.024
3	0.227	0.221	0.088	0.065	0.160	0.275	0.063	0.066
4	0.220	0.267	0.130	0.108	0.252	0.279	0.089	0.113
5	0.214	0.256	0.170	0.156	0.238	0.251	0.116	0.159
10	0.215	0.208	0.349	0.342	0.216	0.210	0.233	0.345
20	0.149	0.139	0.604	0.615	0.167	0.142	0.434	0.619

<sup>a</sup> Standard errors ranged from 0.0029 to 0.0281 for polygenic values and from 0.0 to 0.004 for inbreeding coefficients. <sup>b</sup> The rate of polygenic gain was computed as the difference between the average polygenic values at generations  $t + 1$  and  $t$ .

The extra accumulated response when using genotype information on a major gene of large effect ( $a = 2.0$ ) disappeared several generations after the favourable allele was fixed in both selection schemes ( $G_T$  and  $I_T$ ). This long-term detrimental effect, however, was not a consequence of using the genotype, but rather was due to differences in the rate of inbreeding between the schemes (*table II*). After fixation, the heritability used in  $I_T$  becomes biased upward and this affects the rate of inbreeding and, thereby, the long-term response [9, 17]. With a gene of large effect the bias in the heritability used was large (0.6 versus 0.2) and this led to a substantial reduction in the rate of inbreeding ( $\Delta F \approx 3\%$  with  $I_T$  versus  $\Delta F \approx 5\%$  with  $G_T$ ). Hence, the long-term effect of using the information on the major gene should be evaluated at the generation where the favourable allele is fixed in both selection schemes ( $G_T$  and  $I_T$ ).

In a complete infinitesimal model the correct heritability to be used in the BLUP evaluation is the one from the base population. However, when a major gene is also segregating the use of the initial total heritability in  $I_T$  is debatable as the changes in the gene frequency are not accounted for with BLUP. In order to assess the effect of the heritability in the selection scheme ignoring the gene, a further study was carried out using different choices of heritability in  $I_T$ . *Table III* shows the genetic gain and the change in allele frequency using three different heritabilities: i) the polygenic ( $h_u^2$ ) and ii) the total heritability in the base population ( $h_t^2$ ), and iii) the total heritability updated each generation

using the new gene frequency (i.e.  $h_{t^*}^2 = (\sigma_u^2 + \sigma_v^2) / (\sigma_u^2 + \sigma_v^2 + \sigma_e^2)$ , where  $\sigma_v^2$  is updated each generation using the new  $p$  but  $\sigma_u^2$  remains constant). The use of the polygenic heritability yielded the lowest genetic gain. When using the total heritability, both  $h_t^2$  and  $h_{t^*}^2$  led to very similar patterns in the polygenic gain and in the frequency of the favourable allele.

**Table III.** Total accumulated genetic gain and frequency of the favourable allele ( $p$ ) obtained from truncation selection ignoring genotype information when the BLUP evaluation was carried out using three different heritabilities: the polygenic heritability in the base population ( $h_u^2$ ), the total heritability in the base population ( $h_t^2$ ) and the total heritability updated each generation ( $h_{t^*}^2$ ). The major gene had an additive large effect ( $a = 2.0$ ). The initial  $p$  was 0.15.<sup>a</sup>

	Total accumulated gain			$h_u^2$	$p$ $h_t^2$	$h_{t^*}^2$
	$h_u^2$	$h_t^2$	$h_{t^*}^2$			
1	1.309	1.309	1.309	0.44	0.44	0.44
2	2.670	2.879	2.902	0.75	0.80	0.80
3	3.656	3.778	3.789	0.97	0.99	0.99
4	4.019	4.075	4.072	1.00	1.00	1.00
5	4.269	4.313	4.327	1.00	1.00	1.00
10	5.344	5.409	5.430	1.00	1.00	1.00
20	6.934	7.252	7.077	1.00	1.00	1.00

<sup>a</sup> Standard errors ranged from 0.003 to 0.028 for genetic values and from 0.0 to 0.01 for frequency of the favourable allele.

### 3.2. Optimal selection

Results from the previous section show that with BLUP selection the advantage of using information on a major gene with additive effect can be maintained after the favourable allele is fixed. This advantage disappeared, however, in the long-term due to a higher accumulation of inbreeding. Schemes using or ignoring the genotype information can be objectively compared at the same inbreeding level with optimal selection (which also uses BLUP estimates of breeding values) since maximum possible gains can be obtained under constrained  $\Delta F$ . *Table IV* shows results from optimal selection with  $\Delta F$  restricted to 0.03. This value was approximately that obtained when ignoring the major gene in truncation selection ( $I_T$ ) and was lower than that obtained with  $G_T$  (see *table II*). As intended, with optimal selection, the increase in inbreeding was maintained at the desired constant rate ( $\Delta F \approx 3\%$ ) over generations and consequently the accumulated inbreeding was very similar for both the scheme using the genotype information ( $G_O$ ) and the scheme ignoring the major gene ( $I_O$ ). The optimum number of individuals selected (which was practically constant across generations) was the same for both sexes (i.e. the optimum mating ratio was one) and higher for  $G_O$  ( $N_s \approx N_d \approx 14$ ) than for  $I_O$  ( $N_s \approx N_d \approx 9$ ). This was expected since the heritability used in the BLUP evaluation was lower in  $G_O$  than in  $I_O$  and so more individuals need to be selected with  $G_O$  to keep the rate of inbreeding at the same value.

**Table IV.** Total accumulated genetic gain, frequency of the favourable allele ( $p$ ), rate of polygenic gain and average inbreeding coefficient per generation ( $t$ ) obtained from optimal selection ignoring ( $I_O$ ) or using ( $G_O$ ) genotype information on a major additive gene of large effect ( $a = 2.0$ ). Genetic gains obtained from  $G_O$  are expressed as deviations from those obtained from  $I_O$ . The rate of inbreeding per generation was restricted to 3 %. The initial  $p$  was 0.15.<sup>a</sup>

$t$	Total accumulated gain		$p$		Rate of polygenic gain <sup>b</sup>		Inbreeding	
	$I_O$	$G_O - I_O$	$I_O$	$G_O$	$I_O$	$G_O$	$I_O$	$G_O$
1.	1.815	0.257	0.55	0.65	0.197	0.072	0.000	0.000
2	3.526	0.182	0.94	1.00	0.151	0.232	0.028	0.021
3	3.988	-0.010	1.00	1.00	0.239	0.269	0.050	0.058
4	4.261	-0.016	1.00	1.00	0.269	0.268	0.068	0.085
5	4.519	-0.005	1.00	1.00	0.258	0.269	0.095	0.109
10	5.644	0.090	1.00	1.00	0.215	0.238	0.222	0.236
20	7.518	0.259	1.00	1.00	0.162	0.175	0.426	0.439

<sup>a</sup> Standard errors ranged from 0.008 to 0.026 for total genetic values, from 0.003 to 0.026 for polygenic values, from 0.00 to 0.005 for frequency of favourable allele and from 0.0 to 0.001 for inbreeding coefficients. <sup>b</sup> The rate of polygenic gain was computed as the difference between the average polygenic values at generations  $t + 1$  and  $t$ .

As in the case of truncation selection, the scheme using the genotype information gave more response before the favourable allele was fixed in both schemes. However, in contrast with truncation selection, the advantage of the scheme using the genotype information was not detectable at the time of fixation when the gene had a large effect ( $a = 2.0$ ). In comparison with truncation selection the optimisation procedure led to a faster increase in the frequency of the favourable allele and therefore to a higher difference in polygenic gain between the schemes using and ignoring genotype information when the gene was still segregating (see also *table II*). After fixation, the rate of polygenic gain was higher in  $G_O$  than in  $I_O$  due to a higher accuracy of evaluation. The difference in the polygenic rates in both schemes increased over time and at  $t = 20$  the total accumulated genetic gain was around 3 % higher in  $G_O$  than in  $I_O$ . With an additive gene of moderate effect ( $a = 0.5$ ), optimal selection was not able to maintain, at the time of fixation, the short-term benefit from using the gene (results not shown). Thus, not only with truncation selection but also with optimal selection, the effect of the additive gene needs to be large in order to obtain more gain with  $G$  than with  $I$  at fixation.

Optimal selection always yielded more gain than truncation selection at a fixed rate of inbreeding. A comparison of  $I_O$  with  $I_T$  (see also *table I*) shows that, for instance, at  $t = 20$  the gain was 4 % higher with optimal selection than with truncation selection. The benefit from optimal selection was expected as the numbers of individuals selected and their contributions are optimised for giving maximum gains. When using genotype information the advantage of optimal selection ( $G_O$ ) over truncation selection ( $G_T$ ) with respect to genetic

gain was even greater (around 9 % by generation 20) and the inbreeding was substantially lower. The higher gain with optimal selection was due to the optimisation of the individuals' contributions given the restriction applied on the rate of inbreeding. The restriction on  $\Delta F$  avoided some of the loss in polygenic gain observed with standard truncation.

### 3.3. Efficiency of optimal selection in a mixed inheritance model

Results from *table IV* show that the optimal selection procedure was able to constrain the rate of inbreeding to the desired value at any generation of selection. However, the parameters considered in *table IV* (major gene effect and restriction on  $\Delta F$ ) led to fixation of the favourable allele after very few generations of selection. The frequency of the favourable allele was 0.94 in  $I_O$  and 1.00 in  $G_O$  at generation 2, the first generation with non-zero inbreeding coefficient. After fixation the system works as an infinitesimal model and previous studies have also shown the efficiency of the method for restricting  $\Delta F$  [10].

In order to investigate whether the procedure is able to restrict the rate of inbreeding to the desired value while the major gene is still segregating, a major gene with smaller effect ( $a = 0.5$ ) and a more severe restriction on  $\Delta F$  ( $\Delta F = 0.5\%$ ) was considered. In theory, more severe restrictions on  $\Delta F$  would lead to an increase in the numbers of individuals to be selected and this together with the smaller effect of the major gene would retard the fixation of the favourable allele. Results for  $a = 0.5$  and  $\Delta F = 0.5\%$  are shown in *table V*. The optimal selection procedure was able to maintain the rate of inbreeding to the desired value before fixation both in the scheme using the genotype information and in the scheme ignoring the genotype information.

*Figure 1* shows a comparison of genetic responses obtained with truncation and optimal selection with  $a = 1.0$ . With truncation selection the rate of inbreeding when using or ignoring the genotype information was 5 and 4 %, respectively. With optimal selection the rate of inbreeding was restricted to the lowest value (4 %). In the short-term there were benefits from using the major gene information (at  $t = 2$  the gain was around 30 % higher when using the gene than when ignoring the gene) and from optimal selection (at  $t = 2$  the gain was around 25 % higher with optimal selection than with truncation selection). The combination of the use of the gene and the optimisation led to an increase in gain of 64 %. In the long-term there was not much difference between using or ignoring the genotype information. However, there was still a benefit from using optimal selection (at  $t = 20$  the gain was around 10 % higher with optimal selection than with truncation selection).

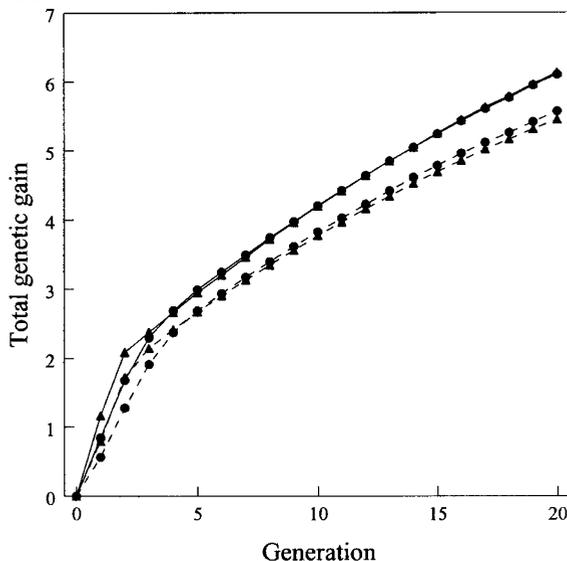
### 3.4. Comparison of BLUP with mass selection

The advantage of the method using the genotype information in mass selection programmes has been previously described for the case when the favourable allele is recessive with large effect (e.g. [15]). *Figure 2* shows a comparison of BLUP and mass truncation selection in this situation. Schemes ignoring the genotype information led to the loss of the favourable allele in some replicates in both mass and BLUP selection. These replicates were excluded from the

**Table V.** Rate of inbreeding ( $\Delta F$ , in %) and frequency of the favourable allele ( $p$ ) per generation ( $t$ ) obtained from optimal selection ignoring ( $I_O$ ) or using ( $G_O$ ) genotype information on a major additive gene of effect  $a = 0.5$ . The initial  $p$  was 0.15. The rate of inbreeding per generation was restricted to 0.5 %.<sup>a</sup>

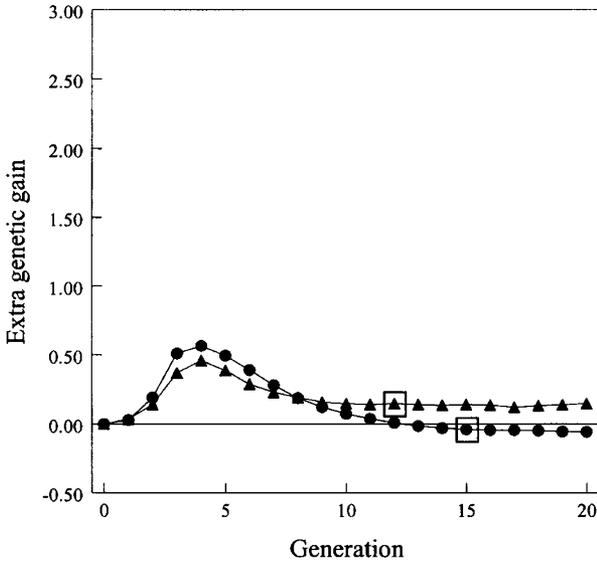
$t$	$I_O$		$G_O$	
	$\Delta F$	$p$	$\Delta F$	$p$
2	0.39	0.23	0.40	0.40
3	0.50	0.29	0.45	0.55
4	0.43	0.35	0.42	0.71
5	0.45	0.42	0.48	0.85
6	0.57	0.49	0.53	0.95
8	0.53	0.64	0.47	1.00
12	0.46	0.87	0.50	1.00
16	0.53	0.96	0.48	1.00
20	0.49	0.99	0.52	1.00

<sup>a</sup> Standard errors ranged from 0.0 to 0.04 for  $\Delta F$  and from 0.0 to 0.005 for frequency of favourable allele.

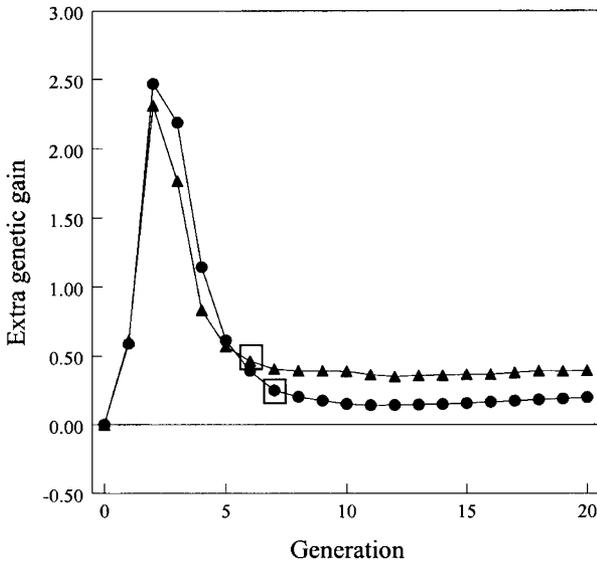


**Figure 1.** Total accumulated genetic gain from truncation (dotted line) and optimal (solid line) BLUP selection when using ( $\blacktriangle$ ) or ignoring ( $\bullet$ ) the genotype information on a major additive gene of effect  $a = 1.0$ . With optimal selection the rate of inbreeding was restricted to 4 % per generation.

analysis. With  $a = -d = 0.5$ , the number of replicates excluded were 58 (out of 500) and 40 (out of 200) in mass and BLUP selection, respectively. There were proportionately more losses with BLUP than with mass selection (11.6 % cf. 20.0 %;  $P < 0.01$ ). The corresponding figures for  $a = -d = 2.0$  were 13 (out of 500) and 5 (out of 200), but this difference was not statistically significant (2.6 % cf. 2.5 %).



(a)



(b)

**Figure 2.** Extra total gain from using the genotype information on a major recessive gene of effect  $a = -d = 0.5$  (a) or  $a = -d = 2.0$  (b) with mass (●) and BLUP (▲) truncation selection. Results are expressed as deviations from the gains obtained when ignoring the genotype information. Squares indicate the generation numbers where the frequency of the favourable allele is greater or equal to 0.95 in both the scheme using genotype information and the scheme ignoring that information.

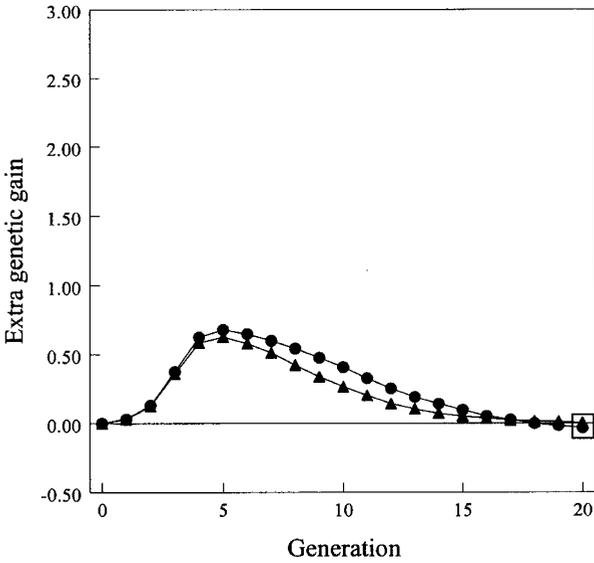
Contrary to the case of additivity of the major locus, the advantage of the method using the genotype information was maintained with BLUP after the favourable allele was fixed when the gene had a moderate effect ( $a = 0.5$ ). With mass selection and  $a = 0.5$  the benefit from using the genotype was lost at fixation. With a gene of larger effect ( $a = 2.0$ ) greater extra gain was obtained in both mass and BLUP selection and the benefit of using the gene information was retained at fixation with both selection methods. In both mass and BLUP selection there was a clear advantage of using the major gene after fixation of the favourable allele although the advantage was higher with BLUP. The maximum extra total gains from using the gene were at  $t = 4$  and  $t = 2$  for  $a = 0.5$  and  $a = 2.0$ , respectively. At the time of this maximum, although BLUP always produces higher gains than mass selection, the extra benefit from using the genotype information was higher with mass selection due to a higher difference in  $p$  between the methods using and ignoring the genotype information.

*Figure 3* shows equivalent results for the case of optimal selection when restricting the rate of inbreeding to 1 % in both mass and BLUP selection. The trends were similar to the case of truncation selection (*figure 2*) although the absolute benefit from using the major gene information was higher before fixation and lower after fixation when optimal selection was applied.

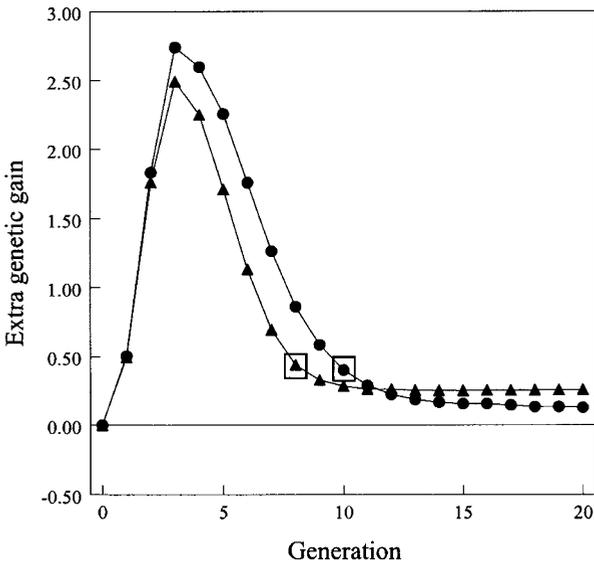
#### 4. DISCUSSION

This study has shown that the apparent conflict between short- and long-term gains reported when the major gene information is present in the genetic evaluation can be made negligible when using a selection tool involving BLUP evaluation and optimisation of selection decisions to maximise response while controlling the rate of inbreeding. Whereas with truncation mass selection it had been reported [5, 6, 12, 15] that selection strategies which explicitly use (rather than ignore) the major gene information to enhance the short-term gain appeared to suffer long-term loss of response, the use of the gene information with the selection tool allowed short-term benefits to be obtained and retained in the long-term. The two components of this tool (BLUP and optimising contributions) both act to counteract the conflict, although perhaps the major impact arises from the optimisation of the genetic contributions of the ancestors. It is notable that the selection tool ignoring the genotype does as well as using the genotype without the selection tool in the short-term and better in the long-term (*figure 1*).

When comparing methods which use genotype information (G) with methods that ignore that information (I) it is useful to divide the selection process into three stages: 1) where the gene is segregating in both G and I; 2) where the gene has been fixed in G but not in I; and 3) where the gene is fixed in both. In the first stage G gives higher total gain owing to a greater accuracy and a greater increase in the frequency of the favourable allele. However, at this stage G gives a lower rate of polygenic gain due to the differential pressure applied to the three genotype classes for the major gene (AA, AB and BB) which leads to a decrease in overall selection intensity applied to the polygenes [15]. In the second stage, G gives higher rate of polygenic gain (which reaches the maximum at this point) but this is the only gain obtained with G whereas I



(a)



(b)

**Figure 3.** Extra total gain from using the genotype information on a major recessive gene of effect  $a = -d = 0.5$  (a) or  $a = -d = 2.0$  (b) with mass (●) and BLUP (▲) optimal selection. Results are expressed as deviations from the gains obtained when ignoring the genotype information. Squares indicate the generation numbers where the frequency of the favourable allele is greater or equal to 0.95 in both the scheme using genotype information and the scheme ignoring that information. The rate of inbreeding was restricted to 1 % per generation.

is still giving gain due to the major gene. In the third stage the comparative gain will depend on the kind of evaluation (e.g. what heritability is used) and this will also affect the rate of inbreeding observed in the two schemes.

The result of using BLUP was that a net benefit at the time of fixation in the I scheme was obtained from using information on the major gene when beneficial alleles were recessive or additive with large effect, but not when additive alleles had small effect. This represents an advantage over mass selection where only recessive major genes of large effect retained the benefit of using the gene in the long-term (e.g. [15]).

The main reason why with BLUP evaluation the long-term loss observed when using genotype information is avoided (or substantially reduced) compared to ignoring the genotype is an extra bias in the EBVs occurring when the major genotype is ignored. The additional bias comes from the fact that with BLUP, the EBV of an individual is regressed toward its parents performance. This regression is appropriate for a complete infinitesimal model, but it is not appropriate when a major gene is segregating. Although the genetic effect due to the major gene is the same for individuals of the same genotype group, BLUP would adjust their EBVs according to their parental mean and this leads to biased estimates. *Table VI* gives an example of the bias induced when ignoring information on the major gene with mass and BLUP selection (when using the gene there is no bias). The bias ( $EBV-g$ ) was calculated in the offspring of a randomly selected base population (so here there is no problem arising either from the use of an incorrect heritability or from the linkage disequilibrium between the major locus and the polygenes). With mass selection the bias is the same for offspring with the same genotype independent of the genotypes of the parents. Thus, within genotypes, the ranking of the candidates is not changed relative to the ranking obtained when using the information on the major gene. However, with BLUP the bias also differs among candidates with the same genotype (i.e. it depends on the genotype of the parents). This causes additional ranking errors within genotypes which affects the polygenic gain achieved.

There are other factors which can also contribute to explain the long-term advantage of using the genotype information with BLUP evaluation. First, the greater accuracy of the polygenic EBVs when using BLUP leads to a reduction in the weight given to the major gene relative to the polygenes, resulting in a greater intensity of selection applied to the polygenes and thus reducing the potential long-term loss. Second, the linkage disequilibrium between the major locus and the polygenic effects induced by selection is expected to be better accounted for when the genotype information is used in the selection criteria. The BLUP evaluation used to estimate the polygenic EBV was carried out on the phenotypic records corrected for the major gene effect. Therefore, it would be expected that any bias in the EBVs as a consequence of the linkage disequilibrium would be substantially reduced. Third, the use of the genotype information to correct the phenotypic records eliminates the problem of bias in the heritability used in the BLUP evaluation. In a complete infinitesimal model the correct heritability to be used is the heritability in the base population. However, when a major gene is segregating, the standard BLUP evaluation does not account for the change in gene frequency due to selection. There is not an appropriate heritability to be used in standard BLUP evaluation when

**Table VI.** Bias in the estimation of breeding values when ignoring the genotype information for offspring genotypes coming from different parental genotypes for a major additive gene ( $a = 2.0$ ) with mass and BLUP selection. The initial frequency of the favourable allele was 0.5.<sup>a</sup>

Offspring	Parents	Mass selection	BLUP selection
AA	AA, AA	-0.535	-0.140
	AA, AB	-0.534	-0.451
	AB, AB	-0.535	-0.762
AB	AA, BB	0.000	0.000
	AA, AB	-0.002	0.311
	BB, AB	-0.001	0.312
	AB, AB	-0.001	0.000
BB	BB, BB	0.531	0.137
	BB, AB	0.534	0.451
	AB, AB	0.533	0.763

<sup>a</sup> Standard errors were around 0.001 in all cases.

the phenotype includes major gene effects. The problem is, however, avoided when the genotypic information is used to correct the phenotype before the BLUP evaluation.

BLUP selection had a greater chance of losing the favourable recessive allele than mass selection. This negative effect of using BLUP in the survival of rare favourable alleles has been reported by Caballero and Santiago [1] who suggested that it is due to the greater reduction in effective population size with BLUP. In addition, procedures which ignore the major gene information have a greater chance of losing the favourable allele and this will be potentiated with small gene effect and low initial frequency of the favourable allele. In these circumstances the benefit of using information on the major gene in the selection criterion could be greatly enhanced.

The most important reason for higher rates of inbreeding observed when using genotype information with BLUP in truncation schemes is the use of a lower heritability (polygenic) relative to that used when ignoring the genotype information (total). Another reason is that the number of favourable alleles the parent carries is a substantial selective advantage for obtaining selected offspring and so between family selection is increased, and consequently so is inbreeding.

The biased heritability used when ignoring the genotype information and the consequent decrease in the rate of inbreeding [9, 17] complicates interpretation in the very long-term. An alternative selection procedure would be to change the heritability used when ignoring the major gene to its polygenic value after fixation. The benefit of the method using the gene would then be retained in the long-term. Selection in two stages (with an initial within family selection on the major gene and subsequent selection on the polygene and the major gene) could also increase the benefits from using genotype information and BLUP although these benefits are small in the long-term [7, 8].

Comparisons of rate of gains in truncation selection schemes were made in the context of static schemes and hence with different rates of inbreeding. It is natural to compare the schemes at the same rates of inbreeding, and when selection was carried out using the dynamic selection tool of Grundy et al. [10] the short-term benefit of using the major genotype was still retained. Moreover, the optimal selection procedure eliminated, or at least substantially reduced, the long-term loss often observed in truncation selection schemes when using the genotype information. Dynamic selection schemes result in more gain than truncation schemes either with or without the use of genotype information. Therefore, dynamic schemes using BLUP have even less conflict between the long- and the short-term gains than BLUP truncation schemes when compared at the same rate of inbreeding. The use of the selection tool allows both the selection and the mating proportions to be made in relation to the desired expected long-term genetic contribution conditional on all the current information, including the major genotype [10]. It might be expected that similar results would be obtained using the procedure of Meuwissen [14] although this has yet to be confirmed.

The optimality of the tool of Grundy et al. [10] for maximising progress with a constraint on inbreeding has been shown for the infinitesimal model but it should also prove near optimal in the mixed inheritance model. The selection decisions and mating proportions conditional upon the estimated breeding values are independent of the inheritance model. Since in this study the effects of the major genotype are assumed to be known, subtraction from the phenotype and prediction of breeding values using the base polygenic heritability derives true BLUP values. However, the very small advantage from ignoring the gene at fixation (see *table IV*) suggests that there remains an additional multiple-generational problem arising from the partition of the population caused by the major gene. This problem has been addressed by Dekkers and Van Arendonk [2] and by Manfredi et al. [13] who describe procedures for optimising the weight given to the major gene to maximise gain after a given number of generations. Future development to combine elements described in these methods with the operational tool of Grundy et al. [10] may result in maximum gains in both the short- and the long-term.

The methodology has only been applied for the case with identified genes of known effect but it might be anticipated that it would have some benefits to marker-assisted selection (MAS). However, in MAS neither the frequency of the gene of large effect, its magnitude or its recombination events with the marker would be known with certainty at any stage and so the procedure is unlikely to be optimal. Nevertheless, these problems are inherent in the methods of Fernando and Grossman [4] (used for instance by Ruane and Colleau [16]) and it would be anticipated that the selection tool of Grundy et al. [10] would be equally applicable to breeding values estimated using markers and these techniques.

Since the work of Gibson [6] there has been concerns over the conflict between long- and short-term benefits of using known genes in selection schemes. This study has shown that when all the information is used with BLUP evaluations, in static and dynamic schemes with constraints on rates of inbreeding, this conflict is very largely removed and both long- and short-term benefits can be obtained over a wide range of cases. This result provides an

encouraging framework upon which to develop further enhancements such as those considered by Dekkers and Van Arendonk [2].

## ACKNOWLEDGEMENTS

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and by the European Union. SAC also receives financial support from the Scottish Office Agriculture and Fisheries Department. Work in Roslin Institute receives support from the Ministry of Agriculture, Fisheries and Food (UK). We thank Dr G. Simm and Dr L. Gomez-Raya for useful comments.

## REFERENCES

- [1] Caballero A., Santiago E., Survival rates of major genes in selection programmes, in: Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 11–16 January, Armidale, vol. 26, University of New England, Armidale, Australia, 1998, pp. 5–12.
- [2] Dekkers J.C.M., van Arendonk J.A.M., Optimizing selection for quantitative traits with information on an identified locus in outbred populations, *Genet. Res., Camb.* 71 (1998) 257–275.
- [3] Falconer D.S., Mackay T.F.C., *Introduction to Quantitative Genetics*, 4th ed., Longman, 1996.
- [4] Fernando R.L, Grossman M., Marker assisted selection using best linear unbiased prediction, *Genet. Sel. Evol.* 21 (1989) 467–477.
- [5] Fournet F., Elsen J.M., Barbieri M.E., Manfredi E., Effect of including major gene information in mass selection: a stochastic simulation in a small population, *Genet. Sel. Evol.* 29 (1997) 35–56.
- [6] Gibson J.P., Short term gain at the expense of long term response with selection on identified loci, in: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 7–12 August, Guelph, vol. 21, University of Guelph, Guelph, Ontario, Canada, 1994, pp. 201–204.
- [7] Gomez-Raya L., Klemetsdal G., Two-stage selection strategies utilizing marker-QTL information and individual performance, *J. Anim. Sci.* (1999) (in press).
- [8] Gomez-Raya L., Klemetsdal G., Hoeschele I., Two-stage selection strategies utilizing marker-QTL information and individual performance, in: 46th Annual Meeting of the EAAP, Prague, 4–7 September, 1995, p. 44.
- [9] Grundy B., Caballero A., Santiago E., Hill W.G., A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes, *Anim. Prod.* 59 (1994) 465–468.
- [10] Grundy B., Villanueva B., Woolliams J.A., Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development, *Genet. Res., Camb.* 72 (1998) 159–168.
- [11] Hospital F., Moreau L., Lacoudre F., Charcosset A., Gallais A., More on the efficiency of marker-assisted selection, *Theor. Appl. Genet.* 95 (1997) 1181–1189.
- [12] Larzul C., Manfredi E., Elsen J.M., Potential gain from including major gene information in breeding value estimation, *Genet. Sel. Evol.* 29 (1997) 161–184.
- [13] Manfredi E., Barbieri M., Fournet F., Elsen J.M., A dynamic deterministic model to evaluate breeding strategies under mixed inheritance, *Genet. Sel. Evol.* 30 (1998) 127–148.

[14] Meuwissen T.H.E., Maximizing the response of selection with a predefined rate of inbreeding, *J. Anim. Sci.* 75 (1997) 934–940.

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

[16] Ruane J., Colleau J.J., Marker assisted selection for genetic improvement of animal populations when a single QTL is marked, *Genet. Res., Camb.* 66 (1995) 71–83.

[17] Villanueva B., Woolliams J.A., Simm G., Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle, *Genet. Sel. Evol.* 26 (1994) 517–535.

[18] Woolliams J.A., Pong-Wong R., Short- versus long-term responses in breeding schemes, in: 46th Annual Meeting of the EAAP, Prague, 4–7 September, 1995, pp. 35.

[19] Wray N.R., Thompson R., Prediction of rates of inbreeding in selected populations, *Genet. Res., Camb.* 55 (1990) 41–54.

### APPENDIX: Optimal solutions for maximising genetic gain while restricting the rate of inbreeding to a specific value

The optimal solutions are found by maximising the function

$$H_t = \mathbf{c}_t^T \mathbf{EBV}_t - \lambda_0 (\mathbf{c}_t^T \mathbf{A}_t^* \mathbf{c}_t - C_t) - (\mathbf{c}_t^T \mathbf{Q} - \mathbf{h}) \boldsymbol{\lambda} \quad (1)$$

where  $\mathbf{c}_t$  is the vector of mating proportions of the  $N$  selection candidates at generation  $t$  (i.e. genetic contributions of the selection candidates to the next generation),  $\mathbf{EBV}$  is the vector of estimated breeding values obtained from BLUP,  $\mathbf{A}^*$  is the modified numerator relationship matrix (augmented  $\mathbf{A}$ ) of candidates [10],  $\mathbf{Q}$  is a known incidence matrix  $N \times 2$  with ones for males and zeros for females in the first column and ones for females and zeros for males in the second column,  $C$  is the constraint on the rate of inbreeding,  $\mathbf{h}$  is a vector of halves of order 2 and  $\lambda_0$  and  $\boldsymbol{\lambda}$  (a vector of order 2) are Lagrangian multipliers. The augmented  $\mathbf{A}$  matrix [10] at generation  $t$  is obtained as

$$\mathbf{A}_t^* = \mathbf{Z}_{t-1} \mathbf{A}_{t-1}^* \mathbf{Z}_{t-1}^T + \mathbf{D}$$

where  $\mathbf{D}$  is a diagonal matrix with elements equal to 1/2 and  $\mathbf{Z}_{t-1}$  is a  $N \times N$  matrix relating individuals of generation  $t - 1$  to generation  $t - 2$  and whose elements are either 0 or 1/2. Element  $(p, q)$  of  $\mathbf{Z}_t$  is 1/2 if the individual  $q$  of generation  $t$  is a parent of individual  $p$  in generation  $t + 1$  [19]. For  $t = 0$ ,  $\mathbf{A}_0^* = \mathbf{I}$ . The constraint used to obtain a constant rate of inbreeding over generations was  $C_t = [\Delta F(1 - 3\Delta F + 12\Delta F^2)]t$ , where  $\Delta F$  is the desired rate of inbreeding [10].

Maximisation of  $H_t$  is equivalent to maximising genetic gain at generation  $t + 1$  ( $\mathbf{c}_t^T \mathbf{EBV}_t$ ) under a constraint on the rate of inbreeding ( $\mathbf{c}_t^T \mathbf{A}_t^* \mathbf{c}_t = C_t$ ) and on mating proportions ( $\mathbf{c}_t^T \mathbf{Q} = \mathbf{h}$ ; i.e. the sum of contributions for each sex adds up to 1/2). Expressions for solving explicitly equation (1) for  $\mathbf{c}_t$  are given by Meuwissen [14]. With this procedure solutions for some animals can be negative ( $c_i < 0$ , for some  $i$ ). As in Meuwissen [14], animals with negative contributions are eliminated from the optimisation which is repeated until all  $c_i$  are non-negative. A contribution  $c_i = 0$  indicates that individual  $i$  is not selected.