Original article

Balancing selection response and inbreeding by including predicted stabilised genetic contributions in selection decisions

JR Brisbane*, JP Gibson

University of Guelph, Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, Guelph, ON N1G 2W1, Canada

Summary – A selection strategy is investigated which should improve upon methodology previously introduced for reducing inbreeding by including genetic relationships in selection decisions. The new strategy includes predictions of stabilised genetic contributions of parents to descendants in selection decisions. An additive infinitesimal genetic model is assumed with discrete generations of selection and random mating of selected parents. Stochastic simulation is used to compare rates of inbreeding and genetic gain from the strategy using relationships with those from the strategy using predicted genetic contributions. The latter strategy gives slightly higher genetic gain at a given level of cumulate inbreeding, but the advantage is small, and the calculations are more complex and difficult to apply in practice, and therefore the previous strategy using relationships is more useful for practical application.

inbreeding / selection / genetic gain

Résumé – Un compromis entre la réponse à la sélection et la consanguinité obtenu en considérant les contributions génétiques à l'équilibre des parents sélectionnés. *Une stratégie de sélection qui devrait améliorer la méthodologie précédemment suggérée pour réduire la consanguinité par l'utilisation des relations génétiques dans les décisions de sélection est étudiée. Cette nouvelle stratégie utilise la prédiction de la contribution génétique à l'équilibre des parents à leurs descendants dans les décisions de sélection. On a supposé un modèle polygénique additif avec des générations discrètes de sélection de même que la panmixie entre les parents sélectionnés. Une modélisation stochastique a été utilisée pour comparer le taux d'accroissement de la consanguinité et le progrès génétique d'une méthode utilisant les relations génétiques à une autre méthode utilisant la prédiction des contributions génétiques. Cette dernière stratégie donne un progrès génétique légèrement*

* Present address: Canadian Centre for Swine Improvement, 2200 Walkley Road, Ottawa, ON K1G 4G8, Canada.

supérieur pour un niveau donné de consanguinité. Cependant cet avantage est faible et les calculs sont plus complexes et plus difficiles à appliquer en pratique. Par conséquent, la stratégie utilisant les relations génétiques s'avère plus utile.

consanguinité / sélection / progrès génétique

INTRODUCTION

In most breeding schemes a balance between genetic gain and inbreeding is sought. Increased genetic gain in the short term is usually associated with increased inbreeding which leads to decreased genetic gain in the long term, due to declines in fitness and genetic variance. Evaluation using the records of all relatives (*eg*, best linear unbiased prediction using an animal model), increased female reproductive rates (*eg*, use of multiple ovulation and embryo transfer or *in vitro* embryo production), and selection of animals at a younger age using pedigree rather than progeny information, lead to increased inbreeding. Various studies (*eg*, Toro and Perez-Enciso, 1990; Verrier *et al*, 1993; Grundy *et al*, 1994; Wray and Goddard, 1994) have investigated selection methods for reducing inbreeding while maintaining high rates of genetic gain. Brisbane and Gibson (1995) showed that a selection strategy (using adjusted estimated breeding value and denoted ADJEBV) that includes genetic relationships in selection decisions gives greater genetic gain at a given level of cumulated inbreeding than selection on a family index with reduced weight on sib information, selection on an index omitting some sib information, or selection on an index with a restriction on the number of full-sibs selected. The objective of this study is to investigate the extent to which the ADJEBV method can be improved by including a prediction of the stabilised genetic contributions of selected animals in the prediction of their effect on inbreeding.

THEORY

The selection objective is assumed to be $M = G_n - D \cdot F_n$, where G_n and F_n are the genetic mean and mean inbreeding coefficient after n generations of selection, and D is the value of a unit of inbreeding relative to a unit of genetic gain. If the genetic contribution of an ancestor is the proportion of genes originating from that ancestor, then with discrete generations, equal family sizes prior to selection, and Nm sires and Nf dams in each generation, each sire has a contribution of $1/(2Nm)$ and each dam a contribution of $1/(2Nf)$ to progeny prior to selection. Contributions of individual sires and dams to the gene pool in subsequent generations vary depending on the genetic merit of the sire or dam and on Mendelian sampling and environmental contributions to the estimated breeding values (EBV) of descendants, although the average contribution remains at $1/(2Nm)$ across sires and $1/(2Nf)$ across dams. The genetic contribution of an ancestor reaches a stable value after a sufficient number of descendant generations. Wray and Thompson (1990) derived analytically that under random mating, the asymptotic rate of inbreeding

is equal to one quarter of the sum of squares of stabilised genetic contributions of any generation of ancestors to descendants.

The mean relationship among animals of any generation is a weighted sum of squares of contributions from all ancestors to parents of that generation (Wray and Thompson, 1990; Brisbane and Gibson, 1995). If it is assumed that contributions of all ancestors to animals in generation t have reached their stabilised values, then the mean relationship, $\bar{a}_n$, among animals in generation n, where $n > t$, is equal to the mean relationship among animals in generation t plus the weighted sum of squares of contributions to parents of generation n from ancestors in generations between t and n. Under this assumption, $\bar{a}_n$ is equal to $\bar{a}_t$ plus a term independent of the selection decision in generation t, and in generation t it seems reasonable to use $\bar{a}_t$ as a predictor of the effect of the selection decision of F_n. The assumption that contributions of ancestors to animals in generation t have reached their stabilised values is not true, but changes in the contributions of ancestors of generation t in subsequent generations will be influenced to a large degree by Mendelian sampling and environmental effects, which are random events. There is a positive linear regression of stabilised contribution on the breeding value of an ancestor (Wray and Thompson, 1990). Given a consistent selection strategy followed in each generation, this regression should enable some prediction of changes in contributions of ancestors of generation t in subsequent generations. The selection strategy proposed here is to use the sum of squares of predicted stabilised contributions of ancestors of generation t as a predictor of the effect of the selection decision of F_n.

Breeding values are not known, but are estimated, and an individual's EBV should be of some use in predicting its stabilised genetic contributions. The usefulness of the EBV will depend on its accuracy. When evaluation is based on an index of the records of collateral relatives, and no pedigree information is used, the covariance between stabilised genetic contribution, $v_{\infty,i}$ and EBV for animal i, at a given true breeding value, A_i, is zero, since prediction errors are not inherited. In this situation, using conditional covariance, and neglecting the effect of selection on the variance of EBV and the genetic variance among parents, we have.

$$\operatorname{cov}(v_{\infty,i}, \operatorname{EBV}_i | A_i) = \operatorname{cov}(v_{\infty,i}, \operatorname{EBV}_i) - \frac{\operatorname{cov}(v_{\infty,i}, A_i) \cdot \operatorname{cov}(\operatorname{EBV}_i, A_i)}{\sigma_A^2}$$
$$\Rightarrow 0 = \operatorname{cov}(v_{\infty,i}, \operatorname{EBV}_i) - r^2 \cdot \operatorname{cov}(v_{\infty,i}, A_i)$$
$$\Rightarrow \operatorname{cov}(v_{\infty,i}, \operatorname{EBV}_i) = r^2 \cdot \operatorname{cov}(v_{\infty,i}, A_i) \tag{1}$$

where r is the accuracy of evaluation and σ_A^2 is the additive genetic variance. It follows that the regression, b, of stabilised genetic contribution on EBV is equal to the regression of stabilised genetic contribution on true breeding value, since

$$b = \frac{\operatorname{cov}(v_{\infty,i}, \operatorname{EBV}_i)}{V(\operatorname{EBV}_i)} = r^2 \cdot \frac{\operatorname{cov}(v_{\infty,i}, A_i)}{r^2 \cdot \sigma_A^2} = b_{v,A} \tag{2}$$

where $b_{v,A}$ is the regression of stabilised genetic contribution on true breeding value. If $v_{\infty,i,*}$ is the predicted stabilised genetic contribution of animal i, based on the EBV, and $v_{\infty,i,**}$ is the predicted stabilised genetic contribution of animal i, based

on the true breeding value, then

$$V(v_{\infty,i,*}) = b^2 V(\text{EBV}_i) = b_{v,A}^2 r^2 \sigma_A^2 = r^2 V(v_{\infty,i,**})$$

Using the EBV, it is possible to account for a proportion r^2 of the variance of stabilised genetic contributions which is associated with breeding value. In the case where genetic evaluation includes pedigree information, prediction errors are inherited to some extent, and $\text{cov}(v_{\infty,i}, \text{EBV}_i | A_i) > 0$. Therefore $\text{cov}(v_{\infty,i}, \text{EBV}_i) > r^2 \text{cov}(v_{\infty,i} A_i)$. This means that $b > b_{v,A}$ and $V(v_{\infty,i,*}) > r^2 V(v_{\infty,i,**})$.

The regression of stabilised genetic contribution on ancestral breeding value is the same for both sexes of descendants, but is different for each sex of ancestors. If b_{xy} denotes the regression of stabilised contribution to descendants of sex y from ancestors of sex x on the breeding value of those ancestors, where $x = m$ (males) and $y = m$ or f, then $b_{mx} = (Nf/Nm) \cdot b_{fy}$ (Wray and Thompson, 1990). b_{mx} will be referred to as b_m and b_{fy} as b_f. In the ADJEBV strategy of Brisbane and Gibson (1995), where contributions of all ancestors are assumed to have reached their stabilised values, v is a column vector with elements 1 to Nm equal to $1/(2Nm)$ and elements $Nm+1$ to $Nm+Nf$ equal to $1/(2Nf)$. The population selection criterion to be maximised is

$$\frac{1}{2}\overline{\text{EBV}}_{s.} + \frac{1}{2}\overline{\text{EBV}}_{d.} - \frac{1}{2}k\mathbf{v}'\mathbf{A}\mathbf{v} = \frac{1}{2}\overline{\text{EBV}}_{s.} + \frac{1}{2}\overline{\text{EBV}}_{d.} - \frac{1}{2}k\left[\frac{1}{4}\overline{a}_{s..} + \frac{1}{2}\overline{a}_{sd..} + \frac{1}{4}\overline{a}_{d..}\right] \quad [3]$$

where $\overline{a}_s$, $\overline{a}_d$ and $\overline{a}_{sd}$ are the mean relationships among selected sires, among selected dams, and between selected sires and dams, $\overline{EBV}_{s.}$ and $\overline{\text{EBV}}_{d.}$ are the mean EBV of selected sires and dams, and k is an arbitrary constant. The selection strategy attempts to maximise this function in each generation. We now replace $\mathbf{v}$ with a vector $\mathbf{v}_{\infty,*}$. Here, and later in this paper, the subscript $*$ is used to denote a prediction based on the EBV. Element i of $\mathbf{v}_{\infty,*}$ is

$$v_{\infty,i,*} = \frac{1}{2Nm} + b_m(\text{EBV}_i - \overline{\text{EBV}}_{s.}), \qquad i = 1, 2, \ldots, Nm, \quad [4]$$

and

$$v_{\infty,i,*} = \frac{1}{2Nf} + b_f(\text{EBV}_i - \overline{\text{EBV}}_{d.}), \qquad i = Nm+1, \ Nm+2, \ldots, Nm+Nf \quad [5]$$

where b_m and b_f are the regressions of stabilised genetic contribution on EBV for sires and dams. The population selection criterion to be maximised when selecting parents in generation t is now

$$\frac{1}{2}\overline{EBV}_{t,s.} + \frac{1}{2}\overline{\text{EBV}}_{t,d.} - \frac{1}{2}k\mathbf{v}'_{\infty,*}\mathbf{A}_{tt}\mathbf{v}_{\infty,*} \quad [6]$$

where $\mathbf{A}_{tt}$ is the relationship matrix among the selected parents and k is an arbitrary constant. $\mathbf{v}'_{\infty,*}\mathbf{A}_{tt}\mathbf{v}_{\infty,*}$ differs from $\mathbf{v}'\mathbf{A}_{tt}\mathbf{v}$ in that relationships involving parents of higher than average EBV, and therefore higher than average predicted stabilised genetic contribution, are given more weight. It can be shown that $\mathbf{v}'_{\infty,*}\mathbf{A}_{tt}\mathbf{v}_{\infty,*}$ is a weighted sum of squares of predicted stabilised contributions of ancestors prior to generation t (Wray and Thompson, 1990; Brisbane, 1994).

METHODS

An additive infinitesimal genetic model, discrete generations of selection, and random mating in a hierarchical design are assumed. Stochastic simulation is used with methodology as given by Brisbane and Gibson (1995). The units of genetic merit are base population genetic standard deviations, $\sigma_{A0} = 1$, and in each generation there are Nm sires, Nf dams, and $n_w/2$ progeny of each sex per dam. The selection method based on predicted stabilised genetic contributions is denoted ADJEBV(R). Both ADJEBV and ADJEBV(R) are simulated, and the balance of inbreeding and genetic gain achieved by each after 6 generations of selection is compared. Parameters of $Nm = Nf = 5$, $n_w = 12$, and $h^2 = 0.5$ are used. A small population size and simple structure are used to minimise the substantial computation involved in the simulation of ADJEBV(R), but results may apply more broadly, since the behaviour of ADJEBV was consistent across a wide range of population sizes and parameters (Brisbane and Gibson, 1995).

ADJEBV

Following Brisbane and Gibson (1995), in each generation Nm sires and Nf dams are initially selected by truncation on EBV based on a family selection index of the individual record and the records of the 11 full sibs. The selected group is then modified as follows. Adjusted EBV are calculated for selected and unselected males as

$$\mathrm{EBV}_{s,i,adj} = \frac{1}{2}\mathrm{EBV}_{s,i.} - k\left[\frac{1}{8}\overline{a}_{s,i.} + \frac{1}{4}\overline{a}_{sd,i.}\right] \qquad [7]$$

where $\overline{a}_{s,i.}$ and $\overline{a}_{sd,i.}$ are the mean relationships of male i with selected sires and with selected dams, and $\mathrm{EBV}_{s,i.}$ is the EBV of male i. Adjusted EBV are calculated analogously for females, and mean relationships are calculated in such a way that the relationship of the animal with itself carries the same weight for a selected animal as for an unselected animal. The unselected male with the highest adjusted EBV replaces the selected male with the lowest adjusted EBV. If the population selection criterion given by [3] is increased the switch is accepted and adjusted EBV for animals recalculated to account for the change in the selected group. Switching and updating of adjusted EBV continues, alternating between the sexes until the population selection criterion cannot be increased (see Brisbane and Gibson, 1995, for further details). Maximising the mean adjusted EBV of selected females and of selected males maximises the population selection criterion, but the process does not guarantee finding the selected group which gives this result.

ADJEBV(R)

In each generation Nm sires and Nf dams are selected initially by truncation on EBV based on a family selection index of the individual record and the records of the full sibs. The adjusted EBV of male i is calculated as

$$\text{EBV}_{s,i,adj} = \frac{1}{2}\text{EBV}_{s,i} - \frac{1}{2}k.Nm.w_{\infty,i,*} \cdot \left(q_2 \cdot q_1 \cdot \left[\sum_{j=1,j\neq i}^{Nm} a_{s,ij}v_{\infty,j,*} \right] \right.$$

$$\left. + a_{s,ii}w_{\infty,i,*} + \sum_{j=1}^{Nf} a_{sd,i,j} \cdot v_{\infty,Nm+j,*} \right) \qquad [8]$$

$$i = 1, 2, \ldots \frac{Nf \cdot n_w}{2}$$

where $q_1 = Nm/(Nm-1)$ if male i is currently selected, or 1 if male i is currently unselected, and $q_2 = (Nm-1)/Nm$. q_1 and q_2 are multipliers which are required to obtain fair comparison of selected and unselected animals, accounting for the relationship of the animal with itself as for ADJEBV (Brisbane and Gibson, 1995). w_{∞,i^*} is the predicted stabilised genetic contribution of the ith male selection candidate, calculated from the deviation of its EBV from the mean of those of the selected sires using equation [4]. If the ith male selection candidate is currently selected then $w_{\infty,i,*}$ will appear in the vector $v_{\infty,*}$. $a_{s,ij}$ is the relationship between the ith male selection candidate and the jth currently selected sire, and $a_{sd,ij}$ is the relationship between the ith male selection candidate and the jth currently selected dam. Adjusted EBV of females are obtained by analogy. The process of switching and updating of adjusted EBV then continues as described for ADJEBV, using the population selection criterion given by [6].

Initially b_m and b_f are unknown, because they depend on the selection strategy of which they themselves are to be part. Initially, therefore 2 500 replicates are run with $b_m = b_f = 0$, equivalent to the ADJEBV method, since $v_{\infty,*} = v$. b_m and b_f are calculated retrospectively using the EBV of base sires and dams, and their genetic contributions to progeny in generation 7. In this example, the expected regression of asymptotic genetic contribution is the same in each sex, and so the average of b_m and b_f, b, is taken. The simulations are then repeated with 3 000 replicates using the average estimated value of b, and a new estimate of b is obtained from the resulting generation 7 regressions. This cycle is continued until the average value of b calculated is close to that used in the selection method. The process is repeated for various values of k in equation [8], in order to determine the performance of the strategy in terms of the rate of genetic gain achieved at any level of inbreeding, compared to that of ADJEBV.

RESULTS AND DISCUSSION

Table I shows examples of the realised values of the regression of genetic contribution to generation 7 on EBV for base sires and dams in each cycle of simulation for various values of k. The regressions increase as k decreases, as expected since more emphasis is put on EBV during selection. The regressions move toward convergence after 3 or 4 cycles. After the second cycle of simulation, subsequent changes in the values of the regressions are much smaller than the standard errors. This means that it was not possible to improve convergence beyond the second cycle of this method. To avoid this problem, at a given value of k, the same seed was used to

initialise the random number generator for every cycle of the simulation. Thus each cycle began with the same replicated base populations, and the sequence of regressions converged to a value specific to those populations. The standard errors of the regressions reflect the sampling variance associated with each estimate across base populations. There was a correlation of around -0.24 between realised values of b_m and b_f across replicates, which contributed to a small reduction in the standard error of the mean regressions given in table I.

Table I. Realised regressions of genetic contribution on EBV (mean $\pm$ standard error, based on 2 500 replicate simulations in cycle 1 and 3 000 in later cycles)[a].

Cycle	Sire regression	Dam regression	Mean regression
$k = 3$			
1	0.0563 ± 0.0038	0.0487 ± 0.0040	0.0525 ± 0.0024
2	0.0505 ± 0.0033	0.0471 ± 0.0034	0.0488 ± 0.0021
3	0.0511 ± 0.0033	0.0472 ± 0.0033	0.0492 ± 0.0020
4	0.0511 ± 0.0033	0.0470 ± 0.0033	0.0490 ± 0.0020
5	0.0512 ± 0.0033	0.0473 ± 0.0033	0.0492 ± 0.0020
$k = 7$			
1	0.0259 ± 0.0023	0.0164 ± 0.0024	0.0212 ± 0.0014
2	0.0177 ± 0.0017	0.0172 ± 0.0017	0.0174 ± 0.0010
3	0.0178 ± 0.0017	0.0159 ± 0.0018	0.0168 ± 0.0011
4	0.0178 ± 0.0017	0.0162 ± 0.0018	0.0170 ± 0.0011
5	0.0177 ± 0.0017	0.0164 ± 0.0018	0.0170 ± 0.0011
$k = 13$			
1	0.00114 ± 0.0014	0.0111 ± 0.0014	0.0112 ± 0.0009
2	0.0064 ± 0.0010	0.0056 ± 0.0010	0.0060 ± 0.0006
3	0.0087 ± 0.0011	0.0076 ± 0.0011	0.0082 ± 0.0007
4	0.0076 ± 0.0010	0.0068 ± 0.0011	0.0072 ± 0.0006
5	0.0074 ± 0.0011	0.0066 ± 0.0011	0.0070 ± 0.0006

[a] k is an arbitrary constant; see equation [8].

Figure 1 shows genetic gain plotted against cumulate inbreeding at generation 7 for ADJEBV and ADJEBV(R). The k values used are given with the figure, and for both lines, points at greater cumulate inbreeding values are always obtained using smaller values of k. ADJEBV(R) gives up to 0.03 units more genetic gain than ADJEBV at a given rate of inbreeding. This advantage is small, but statistically significant, since the standard errors of the mean gains are 0.012 to 0.013 with 3 000 replicates used. A large heritability (0.5) was used so that the EBV and the predicted genetic contributions would be more accurate, and the ADJEBV(R) strategy would be compared in a favourable situation. It was shown earlier that, neglecting some effects of selection on variances, a proportion r^2 of the variance of genetic contributions associated with breeding value is associated with EBV.

When the heritability is lower, the variance of the predicted genetic contributions is lower, and ADJEBV(R) becomes more similar to ADJEBV. As heritability approaches zero, predicted contributions approach $1/(2Nm)$ for all sires and $1/(2Nf)$ for all dams, which are the values used by ADJEBV.

Genetic gain (σ_{A0})

Average cumulative inbreeding
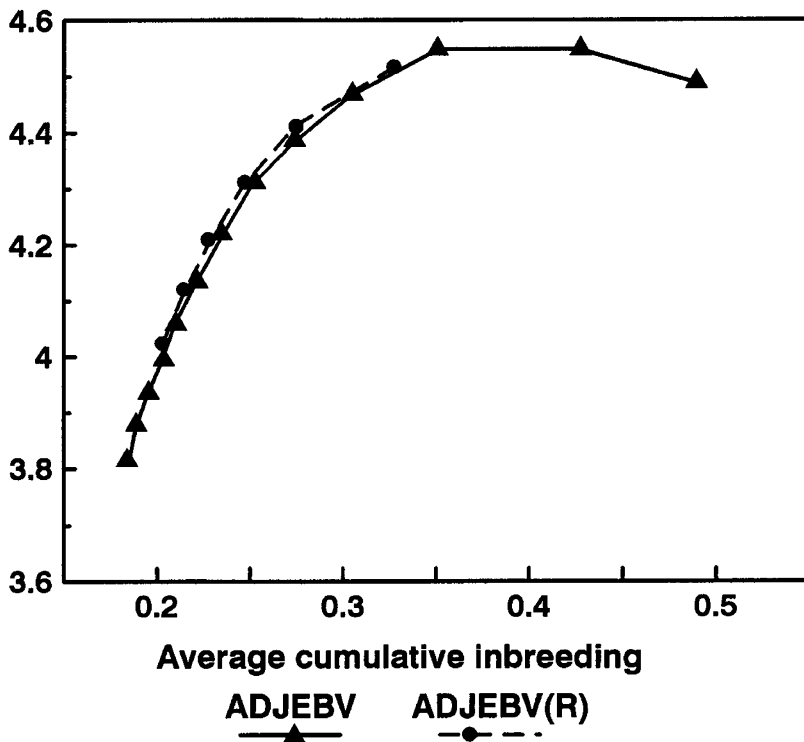
ADJEBV   ADJEBV(R)

Fig 1. Genetic gain *versus* inbreeding for ADJEBV and ADJEBV(R) after 6 generations of selection. $Nm = Nf = 5$, $h^2 = 0.5$, and 6 progeny of each sex per dam. Points are the average of 3 000 replicates and standard errors of the mean genetic gains range from 0.012 to 0.013. Points correspond to k values ranging from 3 to 13 for ADJEBV(R) in equation [8] and from 1 to 23 for ADJEBV in equation [7], in steps of 2 in each case.

These results clearly indicate that failure to include prediction of genetic contributions in method ADJEBV causes trivial loss of performance compared to ADJEBV(R), which is fortuitous given the difficulty in obtaining the appropriate regression. Only a single small population size using single-pair matings was investigated here, and the question of whether the result holds at larger population sizes and different selection intensities is an important one, which needs further study. The advantage of ADJEBV(R) over ADJEBV would be expected to increase with the amount of variation in genetic contributions which can be predicted using ADJEBV(R). Equation [1] used covariances prior to selection. Accounting for the effect of selection, the result in the final line of equation [1] is

$$\text{cov}(v_{\infty,i}, \text{EBV}_i) = \Big(\frac{1 - H}{1 - Hr^2}\Big) r^2 \cdot \text{cov}(v_{\infty,i}, A_i)$$

where H is the proportionate reduction in the variance of the EBV after selection, given by $H = i(i - x)$ where i and x are the selection intensity and truncation point on the standardised normal distribution assuming an infinite population size. Also, the regression coefficient in equation [2] is unaffected by selection. From these results, it follows that using the EBV, we account for a proportion $Q = r^2(1 - H)/(1 - Hr^2)$ of the variance of stabilised genetic contributions which is associated with breeding value. As selection intensity increases, Q decreases from a value r^2 towards zero. However, the regression of stabilised genetic contribution on true breeding value increases with selection intensity (Wray and Thompson, 1990) and so the total variance of stabilised genetic contributions associated with breeding value increases. The overall effect of selection intensity on the advantage of ADJEBV(R) over ADJEBV is therefore not clear.

Further simulation work is required to determine the effects of selection intensity and finite population size on the advantage of ADJEBV(R) over ADJEBV. With overlapping generations, where some animals breed longer and contribute more progeny than others, there will be more variation in the stabilised genetic contributions, and greater potential for ADJEBV(R) to outperform ADJEBV.

REFERENCES

Brisbane JR (1994) Control and prediction of inbreeding in genetic improvement schemes for livestock. PhD Thesis, University of Guelph, Guelph, ON

Brisbane JR, Gibson JP (1995) Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theor Appl Gen* 91, 421-431

Grundy B, Caballero A, Santiago E, Hill WG (1994) A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection progammes *Anim Prod* 59, 465-468

Toro M, Perez-Enciso M (1990) Optimisation of response under restricted inbreeding. *Genet Sel Evol* 22, 93-105

Verrier E, Colleau JJ, Foulley JL (1993) Long-term effect of selection based on the animal model BLUP in a finite population. *Theor Appl Gen* 87, 446-454

Wray NR, Goddard ME (1994) Increasing long-term response to selection. *Genet Sel Evol* 26, 431-451

Wray NR, Thompson R (1990) Prediction of rates of inbreeding in selected populations. *Genet Res* 55, 41-54