GENETIC VARIABILITY WHEN SELECTING ON THE ANIMAL MODEL BLUP

E. Verrier*, J.J. Colleau**, J.L. Foulley**
* : INA Paris-Grignon, 16 rue Claude Bernard, 75231 PARIS Cedex 05, France
** : INRA, Station de Génétique Quantitative et Appliquée, 78352 JOUY-EN-JOSAS Cedex, France

SUMMARY

Several measurements of genetic variability were investigated through simulation in finite populations undergoing selection based on animal model BLUP evaluations. Modifying these indices by putting less emphasis on pedigree information helped to preserve genetic variability. Considering only inbreeding was shown not to give a comprehensive view of the ability of a selected population to deal with changing selection objectives.

INTRODUCTION

The evolution of genetic variability is now a matter of great concern when selecting within finite populations. The genetic variance of the selected trait is reduced through induction of linkage disequilibrium (Bulmer, 1971). Selection changes the family structure (Lush, 1946; Robertson, 1961), which affects the genetic variance of the selected trait but also of other traits of possible interest in the future. Genetic evaluation procedures for livestock are worldwide based on the animal model BLUP (AM-BLUP). Since all genetic relationships are accounted for, stricly selecting on the AM-BLUP is likely to magnify the effect of selection on family structure, as shown from the observed rate of inbreeding in simulated finite populations by Belonsky and Kennedy (1988), Quinton et al. (1992) and Verrier et al (1993). The aim of this paper is to compare, via Monte Carlo simulation, selection indices derived from AM-BLUP as to their ability to maintain sufficient genetic variability in the future, considering inbreeding coefficient and several other criteria.

SIMULATION PROCESSES AND SELECTION PROCEDURES

Phenotypes of populations of constant size were generated for 30 separate generations. The polygenic component (a) was generated according to the statistical model $a_i \sim \text{NID}\left[\frac{1}{2}(a_s + a_d), \frac{1}{2}\sigma^2_{a_0}(1-\bar{F})\right]$ where $\sigma^2_{a_0}$ was the additive genetic variance in the base population and $\bar{F}$ was the average coefficient of inbreeding of the parents (s and d) of a given animal (i). No fixed environmental factor was considered. Three values of the coefficient of heritability (h^2) were chosen : 0.10, 0.25 and 0.50. An independent neutral locus was also simulated : each of the $2N$ variants present in the base parents was labelled from 1 to $2N$ and random transmission to offspring was simulated. In each generation, N_m males and N_f females were selected out of T candidates in each sex, and mated at random. Female prolificacy was constant. Three population sizes were considered :

Populations of size S : $N = 30$; $N_m = 5$; $N_f = 25$; $T = 50$; $\bar{i} = 1.28$

Populations of size M : $N = 120$; $N_m = 20$; $N_f = 100$; $T = 500$; $\bar{i} = 1.78$

Populations of size L : $N = 300$; $N_m = 50$; $N_f = 250$; $T = 500$; $\bar{i} = 1.28$

where $\bar{i}$ is the intensity of selection averaged over the two sexes and computed as if the populations were infinite. The populations were named S10, M10 and L10 for $h^2 = 0.10$, S25, etc. for $h^2 = 0.25$, S50, etc. for $h^2 = 0.50$.

Mass selection was used as the reference selection method. For other selection procedures, the candidates were evaluated by AM-BLUP. Then, a selection index (I_i) was computed by combining the estimates of the mid-parent value (f_i) and of the mendelian sampling component (m_i) :

$$I_i = \omega \hat{f}_i + \hat{m}_i \text{ , with } \hat{f}_i = \frac{1}{2}(\hat{a}_s + \hat{a}_d) \text{ and } \hat{m}_i = \hat{a}_i - \hat{f}_i$$

where $\hat{a}_i$, $\hat{a}_s$ and $\hat{a}_d$ are the predictors of the candidate ($\hat{a}_i$), its sire ($\hat{a}_s$) and its dam ($\hat{a}_d$) genetic values calculated from AM-BLUP theory. Three values were given to the weight ω : 1 (conventional animal model), 1/2 and zero (selection on within-family deviations).

At each generation, the average observed coefficient of inbreeding (F) was computed. Based on average results, the realized effective size at a given generation was computed from the observed change in F between this generation and the next. The realized effective size for the whole selection process was computed as the harmonic mean of successive realized effective sizes.The proportion of founder genes (generation zero) at the neutral locus still existing (i.e. with at least one copy) was computed in each generation. Additionally, but for populations of size S only, the matrix of gene origin [$2T$ offspring x N founder animals] was computed at each generation. Each

5 generations, the probability vector of gene origin was stored to allow a ranking of the founders according to their contribution to the current gene pool. Based on average results, an effective number of founders was computed as $1/\sum p_i^2$ where p_i was the probability of an actual gene coming from a given founder (i). For each situation, 200 replicates were run for populations of size S and 100 replicates were run for larger populations. The values of the different parameters were also given for pure drift, as a reference. Genetic response and genetic variance of the selected trait (mean and variance over replicates) were given in detail in Verrier et al. (1993).

RESULTS

1) Average coefficient of inbreeding

For all populations, the average coefficient of inbreeding was highly affected by selection with ω=1 or 1/2, as shown for population M25 in Figure 1. As expected, selection on within-family deviations resulted in the lowest increase in F. At a given generation, selection method was found to have a significant effect ($P<.001$) on the variance of F between replicates, the AM-BLUP always leading to the highest variance.

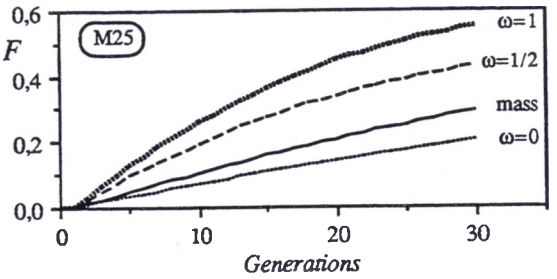


Figure 1. Change in the average coefficient of inbreeding (F) in simulated populations. Mean of 100 replicates.

2) Realized effective population sizes

The realized effective sizes, expressed as a percentage of the effective size for pure drift, are shown for all populations in Table 1. The smaller the initial heritability and the higher the selection intensity, the larger were the differences among selection methods. For the same selection intensity (S and L populations), these differences were all the more important as the number of parents was larger i.e. when pure random drift had less impact. Selecting on the AM-BLUP (ω=1) always led to the largest impact on the realized effective size : in the more unfavourable case (population M10), the realized effective size was equal to only 1/4 of that corresponding to pure drift. On the other hand, the values obtained with ω=0 were close to the expected values under pure drift, so that the choice of families under that selection procedure did not appear to differ from a random process.

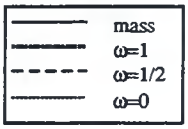
Table 1. Realized effective size of simulated populations, expressed as a percentage of the theoretical value computed for drift with the same numbers of parents

Popu-lation	Pure drift	Mass selection	Selection on BLUP		
			ω = 1	ω = 1/2	ω = 0
S10	16.7	92	45	51	100
S25		87	50	58	102
S50		83	58	70	103
M10	66.7	72	27	35	96
M25		63	28	39	96
M50		57	36	52	95
L10	166.7	81	34	43	99
L25		73	39	51	99
L50		69	49	64	100

3) Loss of founder genes at a given neutral locus

In every population, the loss of founder genes at the neutral locus was dramatic in the first generations, as shown for population M25 in Figure 2. At least 89 % of the founder genes were lost after 10 generations of selection. From generation 10 to generation 30, approximately half to two-thirds of the remaining founder genes were lost (Table 2). Different selection criteria led to different final values, with a maximum loss of genes occuring with selection based on ω=1 or ω=1/2, especially when initial heritability was low and/or selection intensity was high (M populations).

Figure 2.
Proportion of founder neutral genes still existing according to generations (100 replicates).



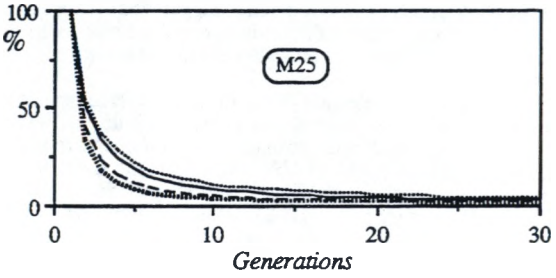


Table 2. Proportion of founder neutral genes (%) still existing after 30 generations of selection. 200 replicates in S populations and 100 replicates in M and L populations.

Population	Pure drift	Mass selection	Selection on BLUP ω = 1	ω = 1/2	ω = 0
S10	4.2	4.0	2.4	2.6	4.3
S25		3.7	2.6	2.9	4.4
S50		3.8	2.8	3.4	4.6
M10	3.5	3.0	1.2	1.5	3.8
M25		2.7	1.3	1.7	3.9
M50		2.5	1.6	2.2	3.7
L10	3.6	3.2	1.5	1.9	3.8
L25		3.0	1.7	2.2	3.7
L50		2.8	2.1	2.6	3.8

4) Loss of founder origins

In S populations, the probability vector of gene origin was the most unbalanced when selecting on AM-BLUP, especially when the initial heritability was low (Table 3). With ω=1, 55 to 66 % of the founders did not contribute at all to the gene pool after 30 generations, according to value of h^2. This loss was very early, as for founder genes loss, but stopped at the 15th generation, contrary to figure 2's results.

Table 3. Effective number of founder animals in the 30th generation (the number of founders in the base population is 30). 200 replicates.

Population	Pure drift	Mass selection	Selection on BLUP ω = 1	ω = 1/2	ω = 0
S10	9.4	8.7	4.8	5.4	9.6
S25		8.2	5.3	6.2	9.8
S50		7.8	6.3	7.4	9.9

DISCUSSION AND CONCLUSION

Using the AM-BLUP as selection criterion clearly led to a great change in family structure in a finite population. The lower the heritability of the selected trait, the larger this change, due to more emphasis on pedigree information resulting in a higher probability of selecting animals from the same lineage. Increasing the selection intensity also induced an increase of the impact of selection based on the AM-BLUP. These conclusions are supported by average results obtained for all parameters considered in this study : average coefficient of inbreeding in a given generation, realized effective size, proportion of founder genes still existing and probability vector of gene origin. Inbreeding increased slowly whereas loss of original genes or founder animals occured abruptly at the first generations. These results suggest that there is a need for a better definition of genetic variability in order to develop relevant optimization procedures. Maintenance of genetic variability is not only a long-term purpose : preserving this variability should deserve much more attention at the beginning of the selection process.

As to the selected trait, selection per se induces a loss of genetic variance through linkage disequilibrium, which is higher when the selection accuracy is higher (Bulmer, 1971). The combined effects of selection on linkage disequilibrium and on family structure led to a decrease in genetic variance which was particularly important when selecting on the AM-BLUP. In some situations where the numbers of parents are very small (e.g. closed selection nuclei), selection methods that preserve genetic variance are able to provide higher overall genetic gains than AM-BLUP (Verrier et al., 1993). In such conditions, procedures putting less emphasis on family information might be interesting alternatives to consider. If other genetic variability measurements are important to consider, one could suggest to preserve variability at the beginning and to relax this constraint after some time. For instance, one could choose in the first generations a low value for the weight assigned to pedigree information (ω in the simulations) and, next, to increase this value from time to time. This needs to be tested via simulation.

The optimal criterion for long-term purposes should additionally take into account the variability of the response. The simulations presented here showed that selection method had a large effect on both mean and variance over replicates of the average coefficient of inbreeding. However, selection method had practically no effect on the variability of response to selection between replicates in seven out the nine simulated populations (Verrier et al., 1993 and unpublished results). This result suggests that, under selection, the link between inbreeding and between-lines genetic variance is not so simple as under pure drift. In order to explain this result, an attempt was made to investigate the variability of the within-replicate genetic variance independently from the average value, by observing the change in the coefficient of variation [$CV(V_A)$] over time, as suggested by Hill (1977). Results in $CV(V_A)$ tended to indicate some differences among selection methods in S populations, but in L populations, the situation was much more complex. Results might be more conclusive if simulation was carried out for a longer selection process. Further research is needed on this topic.

REFERENCES

BELONSKY, G.M. and KENNEDY, B. (1988) J. Anim. Sci., 66 : 1124-1131.

BULMER, M.G. (1971) Am. Nat., 105 : 201-211.

HILL, W.G. (1977) In Proceedings of the International Conference on Quantitative Genetics. eds. Pollak, E., Kempthorne, O. and Bailey, T.B., 343-358.

LUSH, J.L. (1946) Am. Nat., 80 : 318-342.

QUINTON, M., SMITH, C. and GODDARD, M.E. (1992) J. Anim. Sci., 70 : 1060-1067.

ROBERTSON, A. (1961) Genet. Res., 2 : 189-194.

VERRIER, E., COLLEAU, J.J. and FOULLEY, J.L. (1993) Theor. Appl. Genet., 87 : 446-454.