MAXIMIZING GENETIC GAIN IN CONSTRAINED BREEDING SCHEMES

by

J.A. Woolliams

B.A. (Cantab.)
M.A. (Cantab.)

Thesis presented for degree of Doctor of Science
University of Edinburgh

June 1998

CONTENTS

SUMMARY

The thesis records both experimental and theoretical work carried out with the objective of maximizing genetic gain in breeding schemes under a variety of constraints. The experimental work is principally in the context of dairy cattle improvement schemes, whereas the theoretical development is more generally applicable to truncation selection. The work had its origins in the introduction of multiple ovulation and embryo transfer (MOET) in dairy cattle breeding schemes.

A feature of MOET schemes was (i) the low accuracy of selection in juvenile schemes and (ii) the element of random selection among individuals of identical pedigree in both juvenile and adult schemes. An option that offered substantial increases in gain (ΔG) was the prospect of a juvenile predictor and this potential benefit was quantified (Paper 1). It was also noted that a potential predictor had already been identified in that other studies had already observed an apparent genetic association (e.g. Papers 3,8). Subsequent studies, showed this apparent association to be unrepeatable (Paper 4) and, in doing so, established a more stringent testing procedure. The experimental work developed a new successful approach to the problem via growth hormone release (Papers 5,6). The associations were replicated and an understanding of the association in terms of endogenous growth hormone profiles was demonstrated (Paper 7).

The technique of MOET as initially described represented a substantial increase in risk to the breeding scheme as measured by ΔF or variance of response. The element of within-family selection was highlighted as a benefit of a juvenile predictor. However it was clear that there was a dearth of methods to predict inbreeding under selection. This led to the comparison of published and novel 1-generation methods as a means of

predicting ΔF and the identification of their shortcomings as predictors (Paper 9). Further work followed the approach of predicting squared contributions and showed that the method was equivalent to the variance of family size but with a correction for the expected proliferation of lines with selective advantage (Paper 10). The work was applied to mass and sib selection and was shown to yield good, useable predictors (Paper 11).. The approach yielded for the first time expected long-term contributions of an individual in terms of variables conferring selective advantage. A consequence of re-considering the expression of ΔG was that unified equations for ΔG and ΔF were obtained in terms of expectations of functions of long-term contributions, giving a simple commonality of form to these phenomena (Paper 12). The re-expression gave an ideal solution to the problem of maximizing genetic gain with constrained rates of inbreeding in a selection scheme. However whilst the unification proved a useful conceptual approach (e.g. in predicting ΔF in the absence of complete pedigrees) a systematic approach was still lacking. This was overcome by providing expressions for ΔG and ΔF in terms of functions of the expected long-term contributions conditional on variables conferring selective advantage (Papers 13, 32 and 34). These latter terms are derived from two regression models and can be deduced for overlapping generations and for a variety of inheritance modes.

An important principle was established with factorial mating, that inbreeding rates with selection are not simply functions of the numbers of parents, and that the way the parents and the information is used can influence ΔF without a penalty in ΔG or *vice versa* (Paper 14). This principle was extended to applications in new reproductive technologies including nuclear transfer and cloning (Papers 15 and 16) and the results

showed substantial benefits of square designs with future technologies.

An alternative, but related measure of risk in a breeding scheme is the variance of response ($\sigma^2_{\Delta G}$): and, as with ΔF, $\sigma^2_{\Delta G}$ was shown to vary widely with little or no change in ΔG depending upon the design (Paper 17). The methods developed to maximize ΔG with constraints on $\sigma^2_{\Delta G}$ introduced the construction of a quadratic index for selection with differential contributions (which has been further developed in the context of constraining ΔF) and percentile selection. The impact of constraining $\sigma^2_{\Delta G}$ on the design of the breeding scheme was demonstrated in Paper 18. A review of these methods was given in Paper 19 where the connections between constraints imposed on F, ΔF, and $\sigma^2_{\Delta G}$ were explored.

The potential to remove inbreeding from breeding schemes with overlapping generations were investigated using different index forms with simulation (Papers 20, 21, 22, 23 and 24). It was established that considerable inbreeding could be removed when using BLUP evaluation without significant loss of gain. Deterministic models were able to establish optimum designs (numbers of parents, sib-index weights) for fixed resources, risk and time horizons (Papers 25 and 26) in discrete generations. More recently these deterministic models were extended to more general designs to evaluate cloning with constrained ΔF (Paper 27). Application of these principles were used to provide an operational tool to maximize day-to-day gain with constrained inbreeding rates per generation (Paper 33). These used principles such as quadratic indices, introduced previously in the context of constraining the variance of response, and introduced concepts of efficiency based upon the ideal solutions determined in Paper 12.

Other issues related to constrained breeding schemes have been explored. The

incremental inbreeding depression was offset against natural selection to determine target rates of inbreeding from a genetic standpoint (Paper 28). The impact of inbreeding on the overall productivity in sheep was examined using experimental data (Paper 29) to more closely define an economic value for inbreeding: an inbreeding coefficient of 50% was associated with 40% of the outbred productivity. A constraint associated with fitness was examined in the context of the SRY transgenes (Paper 30), where it became clear that the perceived benefits of the transferred gene were negated by its effect in reducing the genetic progress of a breeding scheme. The relationship of genetic contributions to the selection criteria was examined in commercial data in Paper 31.

INDEX OF PAPERS

1. Woolliams, J. A. & Smith, C. (1988) The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46**: 333-345.

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

3. Sinnett-Smith, P. A., Slee, J. & Woolliams, J. A. (1987) Biochemical and physiological responses to metabolic stimuli in Friesian calves of differing genetic merit for milk production. *Animal Production* **44**: 11-19.

4. Woolliams, J. A., Nisbet, R. S. & Løvendahl, P. (1992) The effect of dietary protein on metabolite concentrations during fasting in calves differing genetically in dairy merit. *Animal Production* **54**: 175-181.

5. Løvendahl, P., Woolliams, J. A. & Sinnett-Smith, P. A. (1991) Response of growth hormone to various doses of growth hormone releasing factor and thyrotropin releasing hormone administered separately and in combination to dairy calves. *Canadian Journal of Animal Science* **71**: 1045-1052.

6. Løvendahl, P., Angus, K. D. & Woolliams, J. A. (1991) The effect of genetic selection for milk yield on the response to growth hormone secretagogues in immature cattle. *Journal of Endocrinology* **128**: 419-424.

7. Woolliams, J. A., Angus, K. D. & Wilson, S. B. (1993) Endogenous pulsing and stimulated release of growth hormone in dairy calves of high and low genetic merit. *Animal Production* **56**: 1-8.

8. Woolliams, J. A. & Løvendahl, P. (1991) Physiological attributes of male and

juvenile cattle differing in genetic merit for milk yield: a review. *Livestock Production Science* **29**: 1-16.

9. Wray, N. R., Woolliams, J. A. & Thompson, R. (1990) Methods for predicting rates of inbreeding in selected populations. *Theoretical and Applied Genetics* **80**: 503-512.

10. Woolliams, J. A., Wray, N. R. & Thompson, R. (1993) Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research* **62**: 231-242.

11. Wray, N. R., Woolliams, J. A. & Thompson, R. (1994) Prediction of rates of inbreeding in populations undergoing index selection. *Theoretical and Applied Genetics* **87**: 878-892.

12. Woolliams, J. A. & Thompson, R. (1994) A theory of genetic contributions. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **19**: 127-134.

13. Woolliams, J. A. (1998) A recipe for the design of breeding schemes. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production* **25**: 427-430.

14. Woolliams, J. A. (1989) Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**: 1-14.

15. Woolliams, J. A. (1989) The value of cloning in MOET nucleus breeding schemes for dairy cattle. *Animal Production* **48**: 31-35.

16. Woolliams, J. A. & Wilmut, I. (1989) Embryo manipulation in cattle breeding

and production. *Animal Production* **48**: 3-30.

17. Woolliams, J. A. & Meuwissen, T. H. E. (1993) Decision rules and variance of response in breeding schemes. *Animal Production* **56**: 179-186.

18. Meuwissen, T.H.E. & Woolliams, J.A. (1994) Maximizing genetic response in breeding schemes of dairy cattle with constraints on variance of response. *Journal of Dairy Science* **77**: 1905-1916.

19. Meuwissen, T. H. E. & Woolliams, J. A. (1994) Response versus risk in breeding schemes. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **18**: 236-234.

20. Villanueva, B., Woolliams, J. A. & Simm, G. (1994) Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics Selection Evolution* **26**: 517-535.

21. Villanueva, B., Woolliams, J. A. & Simm, G. (1995) The effect of improved reproductive performance on genetic gain and inbreeding in MOET breeding schemes for beef cattle. *Genetics Selection Evolution* **27**: 347-363.

22. Villanueva, B., Simm, G. & Woolliams, J. A. (1995) Genetic progress and inbreeding for alternative nucleus breeding schemes for beef cattle. *Animal Science* **61**: 231-239.

23. Luo, Z. W., Woolliams, J. A. & Thompson, R. (1995) Controlling inbreeding in dairy MOET nucleus schemes. *Animal Science* **60**: 379-387.

24. Grundy, B., Luo, Z. W., Villanueva, B. & Woolliams, J. A. (1998) The use of Mendelian indices to reduce the rate of inbreeding in selection programmes. *Journal of Animal Breeding and Genetics* **115**: 39-51.

25. Villanueva, B., Woolliams, J. A. & Gjerde, B. (1996) Optimum design for breeding programmes under mass selection with an application in fish breeding. *Animal Science* **63**: 563-576.

26. Villanueva, B. & Woolliams, J. A. (1997) Optimization of breeding programmes under index selection and constrained inbreeding. *Genetical Research* **69**: 145-158.

27. Villanueva, B., Woolliams, J. A. & Simm, G. (1998) Evaluation of embryo sexing and cloning in dairy cattle nucleus schemes under restricted inbreeding. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production* **25**: 451-454.

28. Meuwissen, T. H. E. & Woolliams, J. A. (1994) Effective sizes of livestock populations to prevent a decline in fitness. *Theoretical and Applied Genetics* **89**: 1019-1026.

29. Wiener, G., Lee, G. J. & Woolliams, J. A. (1994) Consequences of inbreeding for financial returns from sheep. *Animal Production* **59**:245-249.

30. Bishop, S. C. & Woolliams, J. A. (1991) Utilization of the sex-determining region Y gene in beef cattle breeding schemes. *Animal Production* **53**: 157-164.

31. Woolliams, J. A. & Mantysaari, E. A. (1995) Genetic contributions of Finnish Ayrshire bulls over four generations. *Animal Science* **61**: 177-187.

32. Woolliams, J. A., Bijma, P. and Villanueva, B. (Unpublished) Expected genetic contributions and their impact on gene flow and genetic gain.

33. Grundy, B., Villanueva, B. and Woolliams, J.A. (Unpublished) Dynamic selection procedures for constrained inbreeding and their consequence for

pedigree development.

34. Woolliams, J.A. (Unpublished) The relationship between rates of inbreeding and expected genetic contributions.

REVIEW OF PAPERS

"And he answered that the beauty of the cosmos derives not only from unity in variety, but also from variety in unity" From "The Name of the Rose" by Umberto Eco.

Paper 1

A feature of MOET schemes was (i) the low accuracy of selection in juvenile schemes and (ii) the element of random selection among individuals of identical pedigree in both juvenile and adult schemes. An option that offered substantial increases in gain (ΔG) was the prospect of a juvenile predictor and this potential benefit was quantified in this paper. These benefits had never been quantified fully and MOET gave new opportunities. However the paper also corrected numerical errors in the paper of Nicholas and Smith (1983), which had originally set the scene for the use of MOET in dairy schemes; the correction reduced the potential benefits of Adult MOET schemes compared to progeny testing, which was the key comparison for implementation.

The paper was important in strengthening the experimental programme on juvenile predictors (Papers 3 to 8). This was my first 'theoretical' 'operational' paper and the work left me with several deep impressions which have been driving forces throughout much of the work that follows. Firstly that the genetic models which were used to guide major investments in breeding schemes were naive and somewhat arbitrary: consideration of inbreeding was based upon untenable assumptions because there was a lack of necessary theory and comparisons of gain were made with many design features (e.g. size, selection intensity) pre-determined. A second impression was that different components of gain had different penalties with respect to inbreeding rates.

Paper 2

This was an invited paper at the 4th World Congress. It develops some of the models used in Paper 1, and others, more rigorously and defines some of the problems tackled later in this thesis.

Paper 3

This study of potential juvenile predictors of dairy merit embodies the general hypotheses underlying most approaches to the problem: that genetic progress had promoted changes in the energy metabolism of cattle when under metabolic challenge. The observations concerning urea were noted in Paper 1 and led to the study described in Paper 4.

Paper 4

This experimental study was indicated by a review of the prior work (including Paper 3 in which the concentration of urea in the blood after a fast appeared to be associated with genetic merit for yield. The result of this study was very disappointing. However the publication of negative results was, and remains, important. One achievement of the study was to raise the standard for the size of physiological trials using selected lines as a means of detecting genetic correlations; the study was more powerful than the sum of the three previous trials that had indicated the association with urea.

Paper 5

This paper was a necessary first step in testing a specific hypothesis that Growth Hormone (GH) release had a positive genetic association with yield. Development of a test which was based on giving GH secretagogues would need knowledge of the response

curves.

Paper 6

Unlike the more general hypotheses of Paper 3, the approach in this study was based upon a specific prediction from work done in lactating cows of differing genetic merit. The doses were chosen to be half-maximal, using Paper 5, in an attempt to harness differences in both sensitivity and maximal response. During the experiment we discovered a paper which had carried out a similar test and appeared to have different results; however, the authors had analysed the data in such a way that they had obscured what was clear from the graphs they themselves had presented, i.e. that high merit calves released more GH.

Paper 7

The result of Paper 6 had to be repeated to have credibility. As well as the potential application in industry there was also considerable scientific interest in the underlying physiology. A natural and achievable extension was to examine the endogenous pulsatility of GH in association with the secretagogue induced release. The confirmation of Paper 6 was very pleasing. The study was only achievable through the experimental skills and dedication of the co-authors.

Paper 8

This is the only uninvited published review I have undertaken.

Paper 9

This paper is one of a series of papers that has attempted to understand inbreeding rate (ΔF) in selected populations. It compares a number of methods that attempt to predict ΔF by considering selection in only one or two generations. An eigenvalue

method was developed by me in Paper 14, another by Burrows (1984), and others considered variances of family sizes as developed by Hill (1979). The work showed that all could be related to each other through the matrix of co-selection probabilities, but all were clearly underestimating the rate of inbreeding.

Paper 10

Whereas half of this thesis has its roots in the paper of Nicholas and Smith on MOET schemes, the other half is firmly rooted in the pioneering paper of Wray and Thompson (1990). The paper of Wray and Thompson sketched a framework for a new approach to predicting inbreeding; they had shown that the rate of inbreeding could be predicted from the sum of squared genetic contributions from a single generation. However the implementation of this approach was recursive with the result that much remained obscure, with complex arguments to predict pathway extension many generations into the future. This paper is also complex (although its published form is much reduced from my first derivation), but it *is* a simplifying paper; it produces a closed form for ΔF, with a simple denouement that the approach of Wray and Thompson was not unfamiliar but was closely related to methods considered in Paper 8 with additional terms that accounted for expected changes in ancestral contributions. As importantly, it showed that expected long-term contributions could be derived as linear functions of variables conferring selective advantage to an ancestor (a point not appreciated by Wray and Thompson, 1990).

Paper 11

This paper achieved an end result of providing practicable predictions of ΔF for sib indices. However the understanding gained from Paper 10 was not sufficient to

produce a streamlined approach to prediction. However much of its complexity is due to the attempt to predict ΔF for classical sib indices, many of which promote high rates of inbreeding (>0.03) with the numbers of parents simulated; these situations demanded special methodologies to cope with the very high index correlations among sibs.

Paper 12

This was an invited paper. Whilst the original intention had been to review the prediction of inbreeding (such as the preceding papers in this thesis) and the simulation work (e.g. Papers 20 to 23), I decided to write down for the first time in a published form what is a general unification of the theories of genetic progress and inbreeding: what are the generating forces in a population and what is the underlying relationship between the forces governing progress (i.e. mean gain) and inbreeding (i.e. variance loss). The relationship is extraordinarily simple and beautiful. Also its form immediately makes connections with work I had been aware of concerning clonal propagation in forestry. Combining these results again resulted in lovely insight into what is ideal in genetic selection.

Paper 13

Once the connections described in Paper 12 are made they become 'obvious' (albeit novel), and described a concept rather than a result, and the value of a concept must lie in its help in generating predictions. The personal time available for the development of the concept was limited with only a very few dedicated periods. Therefore some of the detailed results of how Paper 12 helps prediction and the design and operation of breeding schemes is shown in as yet unpublished work (Papers 32, 33 and 34) and has only been achieved with the involvement of others. Paper 13 gives an

indication of the outcome of these papers. Whilst this was a contributed paper to the 6th World Congress it was given time for presentation over and above that allocated to other contributed papers.

Paper 14

This paper marks the first in a series of papers concerning the design of breeding schemes. Nicholas and Smith (1983) was wide ranging in its treatment of potential technologies for use in breeding schemes. However it did not properly consider constraints and implicitly assumed that there was a one-to-one function between rates of gain and rates of inbreeding within the same resources. This paper showed that this is not the case: there are ways to use genetic variation wisely and, conversely, ways to waste it. The novelty of the paper lies not with the factorial mating *per se*, since forest breeders (for example) had been using this as a design long previously, but in its genetic justification of why factorial mating should be used.

Paper 15

A further problem with the constraints in the paper of Nicholas and Smith (1983) was the tendency to confuse more gain per resource and more resources to produce gain. The treatment of cloning was one such area of confusion. This paper considers cloning within a closed herd with fixed resources and reached a completely different conclusion, since increased cloning reduced selection intensity. It emphasised the importance of proper definition of the problem before deciding upon investment in technology.

Paper 16

This paper was the first invited review to appear in Animal Production (now Animal Science). The invitation and publication marked the centenary of Sir John

Hammond's birth in 1889. Much of what appears on quantitative genetics was novel in trying to explore the use of technologies in constrained systems, even though the methods now appear naive. The principles outlined in the quantitative genetics section remain valid 10 years on and illustrate the ways in which technologies can interact to produce benefits. At the time of writing the manuscript I doubted my co-author's enthusiasm for nuclear transfer compared to some of the other technologies.

Paper 17

This was the first of a series of papers in collaboration with Theo Meuwissen and deals with constraining the variance of response. This constraint was both a tractable and reasonable approach to risk in breeding schemes. In the paper there is an early development of quadratic indices and differential use, later developed by Wray and Goddard (1995) and others in the context of inbreeding. It arrived at a means of describing risk to users of EBVs when faced with a choice between different accuracies i.e. selection based upon percentiles. It seems more obvious and simpler than any explanation given in the catalogues I have seen to date.

Paper 18

This paper gives a full deterministic model to show how dairy cattle breeding schemes change in structure according to the constraint imposed. Dairy cattle breeding schemes are sex- and age-limited and an important decision is the setting up of a large progeny testing scheme to provide bull proofs before widespread use. The results showed that the optimum schemes move from a reliance upon progeny testing to larger schemes using young bulls as the restriction on variance is weakened. The method of solution incorporated simulated annealing which was then a novel approach to the

maximization problem within animal science.

Paper 19

This was the second invited paper I had at the 5th World Congress. It provides a summary of the work described in Papers 17 and 18. It shows the relationship of Paper 17 to the papers of Wray and Goddard (1995) and others where the constraint was on the inbreeding coefficient after a given time rather than on the variance of response.

Paper 20

This examines alternative ways of controlling inbreeding rate in beef breeding schemes through modifying the index of selection. The simulation study was carried out with overlapping generations. The extra degree of freedom the scheme has from varying the generation interval was found to complicate the interpretation. Simulation, chosen in the absence of the necessary theory, yielded at best general conclusions.

Paper 21

An objective of the project which funded much of this work was to evaluate the potential of new technologies to beef breeding schemes when resources are fixed and rates of inbreeding are constrained.

Paper 22

The value of embryo transfer in beef breeding had never been established clearly despite the benefits that were acknowledged in dairy breeding. This paper provides a clear answer with respect to the genetic benefits with realistic parameters for embryo collection and transfer.

Paper 23

Similar to Paper 24, but for dairy breeding schemes.

Paper 24

Paper 12 suggested a range of indices that might be used to maximize gain with constrained inbreeding. These were evaluated in this paper.

Paper 25

The development of deterministic predictions of ΔF for mass selection (e.g. Paper 10) in discrete generations allowed the achievement of flexible optimum design (numbers of parents) determined only by the resources (total number of offspring, time horizon and ΔF). The saving in computer time was immense compared to simulation. The results showed that the optimum mating ratio was not always 1, as thought previously.

Paper 26

This develops the same framework for sib indices (using Paper 11) as Paper 29 did for mass selection. Additional variables for optimization were the index weights. The paper therefore develops the concept of optimum sustainable indices as distinct from the classical indices for maximizing gain in the first generation. There were substantial differences between the classical and optimized weights for reasonable constraints on inbreeding and modest time horizons. Compared to mass selection, optimum sustainable indices only gave substantial extra gain when the schemes had large resources and the heritability was low. This paper, along with Paper 25, begins to answer systematically the general questions that had been avoided previously in breeding scheme design. Thus the problems with Paper 2 were being solved.

Paper 27

Following the publication of Paper 15, Colleau (1992) suggested that by opening the nucleus up and defining resources differently, then cloning may be beneficial for a

breeding scheme. This assessment was carried out assuming mass selection and without formally considering inbreeding rates. Therefore this paper uses the methods of Paper 33 to tackle the problem to examine whether or not cloning can add additional gain to a breeding scheme without increasing rates of inbreeding.

Paper 28

An approach to estimating a desirable effective population size.

Paper 29

Like Paper 20, this examines the value that might be attached to inbreeding rate. It is in fact the culmination of a series of papers (not presented in this thesis) which examines the effect of inbreeding upon different sets of traits relating to primary products, reproduction and survival. This paper estimates the total cost in productivity that accompanies rapid inbreeding. It shows the net loss in overall productivity is very substantial indeed.

Paper 30

This paper arose from the discovery of the sex-determining region of the Y chromosome and demonstrates that the dynamic nature of populations has to be taken into account before deciding upon introducing novel breeding strategies. The results were counter to the beliefs of non-specialists who considered that benefits from super-males were immediate and clear.

Paper 31

An examination of genetic contributions in a real selected population. It attempts to estimate inbreeding rate from differing amounts of pedigree information.

Paper 32 (unpublished)

This paper details one of the results alluded to in Paper 13. It describes the modification of a very basic concept in quantitative genetics, namely gene flow, to account for selection and the consequences this can have for other concepts such as generation interval. It derives fundamental equations for the expected long-term contribution of a selected individual in the form of a linear regression on selective advantages. These methods are then applied to a variety of examples including BLUP, imprinted variation and overlapping generations. It has been submitted to Genetics.

Paper 33 (unpublished)

At one level this paper derives a simple method of maximizing genetic progress whilst giving operational control over rates of inbreeding. The method is easily generalized to overlapping generations. It is explicitly based on controlling long-term genetic contributions. However at a deeper level the paper explores the optimality of the method by showing that it is attempting to deploy selected individuals that in a way that is optimal given the existing information. Further it explicitly derives the upper bound on genetic gain when selecting in a population with given size and rate of inbreeding. This is based on the connections to other work made in Paper 12. It allows the concept of efficiency to be applied to a breeding scheme. This is submitted to Genetical Research.

Paper 34 (unpublished)

This derives the theorem which allows the methodology described in Paper 32 to be applied to rates of inbreeding. It was briefly described in Paper 13.

BIBLIOGRAPHY

Burrows, P.M. (1984) Inbreeding under selection from related families. *Biometrics* **40:** 895-906.

Colleau, J.J. (1992) Combining use of embryo sexing and cloning within mixed MOETs for selection on dairy cattle. *Genetics, Selection, Evolution* **24**: 345-361.

Hill, W.G. (1979) A note on effective population size of populations with overlapping generations. *Genetics* **92**: 278-289.

Nicholas, F.W. and Smith, C. (1983) Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

Wray, N.R. and Thompson, R. (1990) Prediction of rates of inbreeding in selected populations. *Genetical Research* **55:** 41-54.

Wray, N.R. and Goddard, M.E. (1994) Increasing long-term response to selection. *Genetics, Selection, Evolution* **26:** 431-451.

STATEMENT OF AUTHORSHIP

For the papers reported under joint authorship, I have tried to indicate my approximate contribution under the headings: Initiation, including conception and planning; Execution, including the generation of results and any associated statistical analysis; and Completion, including the writing and steering of the paper through to publication.

Paper	Initiation	Execution	C'pletion	Paper	Initiation	Execution	C'pletion
1	90	90	90	18	30	20	20
2	100	100	100	19	40	30	10
3	40	30	30	20	70	10	30
4	100	80	90	21	50	10	30
5	50	40	50	22	30	10	30
6	50	40	50	23	70	10	20
7	100	60	100	24	70	10	70
8	50	70	70	25	40	20	30
9	50	40	40	26	50	20	30
10	100	80	90	27	60	40	30
11	50	30	30	28	20	10	40
12	100	90	90	29	40	10	30
13	100	100	100	30	70	10	50
14	100	100	100	31	90	90	90
15	100	100	100	32	90	80	90
16	50	51	50	33	70	10	30
17	50	60	80	34	100	100	100

None of this work has been submitted for other degrees.

John A. Wallis.

ACKNOWLEDGEMENTS

Firstly I am truly grateful to Dr. G. Wiener (of the Animal Breeding Research Organization) for having the courage to appoint me to a post for which I was barely qualified. From him, from Dr. A. C. Field (of Moredun Research Institute), and from the late Prof. R. B. Land (of the Animal Breeding Research Organization) I learnt much about the importance of livestock breeding, the scientific approach to problems, and a professional approach to science. I am also grateful to Prof. R. B. Heap of the Institute of Animal Physiology and Genetics Research who encouraged me, at critical times, to concentrate on the quality of science. All these influences have combined to provide the philosophical foundation for this thesis; I hope the quality of the thesis is sufficient to justify their encouragement. The solid foundation they provided has been very necessary given the perilous financial climate in which this work has been conducted. The good fellowship of collaborators, colleagues and assistants has been a constant support throughout, especially during the difficult times.

Scientifically the inspiration for this thesis has come from two papers that are landmarks of quantitative genetics: Nicholas and Smith (1983), and Wray and Thompson (1990). Both are in the bibliography but their significance demands a special mention.

Paper 1

Anim. Prod. 1988, 46: 333-345

0003-3561/88/7074033302-00

THE VALUE OF INDICATOR TRAITS IN THE GENETIC IMPROVEMENT OF DAIRY CATTLE

J. A. WOOLLIAMS AND C. SMITH†
AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

ABSTRACT

The value of indicator traits (T), such as physiological or biochemical traits in the genetic improvement of dairy cattle for milk yield (M) was studied. First, some corrections were made to the base rates of genetic change possible by improvement systems based on progeny testing and on multiple ovulation and embryo transfer (MOET), and on combinations of these. Efficient field progeny-testing systems can be competitive with current adult MOET nucleus herd schemes but juvenile MOET nucleus herd schemes offer substantial increases in rates of response. With high co-heritability, selection for the T alone may allow greater rates of response than those currently considered feasible using progeny testing. However, faster rates are obtained with combined selection. When breeding values are accurately measured by pedigree and performance records on M, as in the progeny test, the extra rates of response with combined selection may be small. Where breeding values are less accurately assessed, as in juvenile MOET nucleus schemes, the extra rates of response can be appreciable. For T with co-heritability ($h_M r_G h_T$) of 0·27 and the CV for M from 0·15 to 0·20, response rates of 2·0 to 2·7% of the mean per year possible by traditional methods could be increased to 2·2 to 2·9% in progeny testing schemes, 2·3 to 3·1% and to 4·3 to 5·7% for adult and juvenile MOET nucleus schemes respectively.

A possible useful indicator trait is blood urea nitrogen (BUN) measured in young animals after a short fast. Results from four experiments with calves having high or low genetic merit for M were summarized. The pooled co-heritability estimate was −0·27 (s.e. 0·05). With this, or even a more modest effect, BUN would be a useful indicator trait in selection for milk production. Its use in practice in high and low selection lines or in a section of the industry, would allow assessment of the merit of the method.

INTRODUCTION

THE advent of multiple ovulation and embryo transfer (MOET) has stimulated the re-appraisal of methods of genetic improvement of dairy cattle. A series of papers (Land and Hill, 1975; Nicholas, 1979; Nicholas and Smith, 1983; Christensen, 1984; Colleau, 1985; Christensen and Liboriussen, 1986) has been presented examining the possibilities, with different methods of use of MOET, and has indicated that useful gains in the rates of genetic change are possible. Further advances may come from physiological and other research providing indirect measures (to be called indicator traits) of genetic potential for milk yield, especially if those were available in young animals. These could be used to screen young bulls for progeny testing or, if accurate enough, could replace progeny testing (Wiener, Sinnett-Smith, Slee and Woolliams, 1986). There has been little quantitative assessment of the potential impact of such indicator traits on dairy breeding schemes. This paper develops the work of Walkley and Smith (1980) on indicators for indirect selection (for reproduction in sheep) in the context of dairy breeding schemes, assesses the potential of a variety of possible schemes, and considers the value of a specific, possibly useful, indicator trait in practice.

As a benchmark, the rate of progress with an efficient progeny testing scheme is used. The opportunity is also taken to correct

† Present address: Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1.

errors in the previous results for MOET schemes published by Nicholas and Smith (1983).

METHODS

Assumptions of genetic parameters and selection intensities

Throughout the paper the heritability of milk yield assessed by a single lactation (h^2_M) is assumed to be 0·25, irrespective of lactation number, and the genetic correlation of any pair of lactations has been assumed to be 1. The repeatability of milk yield across lactations has been taken to be 0·35.

The selection intensities used assume large population sizes for both progeny testing and MOET nucleus schemes, in all circumstances except when within-family selection is considered. Any reduction in selection intensity due to positive correlations among relatives in indices used for selection (e.g. in MOET) has not been allowed for. However, the increase in selection intensity above that for normal independent deviates due to the negative correlations among deviations from the observed family mean (when selecting within families) has been calculated using the formula given by Owen and Steck (1962). Thus $i = i_n(n/(n-1))^{1/2}$ where i_n is the selection intensity achieved by selecting one from n independent, normally distributed, random variables of unit variance and i is the selection intensity used.

Rate of change by progeny testing

Possible rates of genetic change by progeny testing for milk yield have usually been set at about 0·10 standard deviation units per year (1·5% or 2% of the mean per year for coefficients of variation of 0·15 and 0·20 respectively) (e.g. Nicholas and Smith, 1983). Brascamp (1978) showed that this base rate can be improved by two feasible changes in the evaluation procedure. These affect the rates of genetic change in the bull breeding nucleus herds. First, cows in nucleus herds could all be mated to the best bulls, the bulls used to breed the next generation of young bulls (the bb path), and all progeny testing of young bulls would take place in testing herds. Thus, the genetic merit of bulls to breed cows (bc) in nucleus herds should be as high as that of bulls to breed bulls (bb), since several (bb) bulls must be used to reduce inbreeding. However, the (bb) path may be shorter than the (bc) path, for young bulls can be produced over a short period while young cows are produced throughout the year. Secondly, cows could be selected to breed bulls after their first lactation, the most reliable for estimating genetic merit, rather than waiting for three lactation records as has been usual in the past. Thus, their sons would be born at their third calving. With these simple changes, the possible rates of genetic change from progeny testing can be increased from 0·104 (Nicholas and Smith, 1983) to 0·133 standard deviation units per year, as shown in Table 1. Selection on estimated breeding value over several age groups, rather than strictly within age groups, might add to this response rate. In practice, the reliability of the records of the cows selected to breed bulls (cb path) may be uncertain due to differential treatments, and this may restrict improvement through this path.

Multiple ovulation and embryo transfer (MOET)

The genetic responses in milk yield possible by use of MOET were derived by Nicholas and Smith (1983) for a range of MOET rates and mating ratios. They considered juvenile MOET nucleus schemes, where selected males and females are bred by MOET at 12 to 14 months of age, giving a generation interval of 22 months, 1·83 years, and adult MOET nucleus schemes with selected males and females bred by MOET at 34 to 36 months of age, giving a generation interval of 44 months, 3·67 years. An error in the derivation was found later (the number of sibs used was twice that stated). To obtain the correct response rates, the figures given in Tables 4, 5, 6 and 7 of Nicholas and Smith (1983) should be reduced by about 1 to 8% (average 3%) of tabulated values. In evaluation of juvenile MOET nucleus, Nicholas and Smith (1983) did not include information on the paternal side of the pedigree, the records on the sire's full-sibs, half-sibs and dam (e.g. Avalos and Smith,

TABLE 1
Possible rate of genetic change per year in milk yield by progeny testing

Path	Age at birth of progeny (years) Range	Average	Accuracy of selection	Proportion selected (s.d. units)	Selection differential (phenotypic s.d. units)	Genetic superiority (phenotypic s.d. units)
Bull-bull(bb)	6·3 to 6·8	6·5	0·88	3/100	2·27	1·00
Bull-cow(bc)	6·3 to 7·8	7·0	0·88	3/100	2·27	1·00
Cow-bull(cb)	4·0 to 4·5	4·3	0·66	1/100	2·66	0·88
Cow-cow(cc)	2·0 to 9·0	4·5	0·60	85/100	0·27	0·08

$$\text{Possible annual response} = \frac{1{\cdot}00 + 1{\cdot}00 + 0{\cdot}88 + 0{\cdot}08}{6{\cdot}5 + 7{\cdot}0 + 4{\cdot}3 + 4{\cdot}5} = 0{\cdot}133 \text{ phenotypic s.d. units.}$$

Selection based on:
bulls: 50 effective daughters;
cows: sire, maternal grandsire, cow (cb, one lactation; cc, zero to eight lactations).

1987). Including this will add to the accuracy of selection, increasing responses in juvenile MOET nucleus schemes by 25 to 35% of the tabulated values in Nicholas and Smith (1983), as shown here in Table 2. These changes increase the superiority of juvenile MOET nucleus schemes over adult MOET nucleus schemes.

In the evaluation here with MOET, two sets of levels for MOET sibship size and mating ratio have been used, a normal (N) rate (four male and four female progeny at selection per full-sibship, eight donors per male) and a high (H) rate (eight males, eight females, and 16 donors). As in Nicholas and Smith (1983), to restrict inbreeding rates, only one male per selected full-sibship is used, but all the females if required. This gives the proportions selected as 1/8 in males and 1/4 in females for the normal MOET rate and 1/16 and 1/8 for the high MOET rate. The rates of response possible range from 0·17 (N) to 0·23 (H) s.d. units per year with juvenile MOET nucleus schemes, and from 0·12 (N) to 0·16 (H) s.d. units per year for adult MOET nucleus schemes, as shown in Table 2. Thus, it takes effective adult MOET schemes to yield higher responses than efficient progeny testing systems. So it is apparent that for appreciable gains in response, juvenile MOET nucleus schemes must be used.

Combining progeny testing of sires with MOET of females has been proposed by Nicholas (1979), Christensen (1984) and Colleau (1985). With efficient progeny testing, as specified in Table 1, MOET in females adds only moderately to response, as shown in Table 2. However, it offers alternative breeding systems, and may avoid the unreliability of the (cb) path if the females are kept in controlled nucleus herds (Christensen, 1984).

Indicator traits

Indicator traits are defined as traits which give indirect information on milk yield, and may be used to predict genetic merit for yield. For example they may be physiological traits related to or involved in milk production. These are assumed to be measured in both sexes before reproductive age. The value of an indicator trait in increasing selection response will depend largely on the magnitude of the co-heritability, the standardized genetic covariance between milk yield (M) and the indicator trait (T), which is measured by $h_M r_G h_T$, where r_G is the genetic correlation of M and T, and h_M^2 and h_T^2 are the respective heritabilities. The genetic correlation is fixed but the effective heritability of the indicator trait can be increased if several measurements are taken, by a factor $[n/(1+(n-1)t)]$ where n is the number of measurements and t is the

TABLE 2

Possible annual genetic change (in phenotypic s.d. units × 1000) in milk yield (M)*, using an indicator trait* (T) *and combined selection, for different selection systems*

		Indicator trait										
Genetic correlation	(r_G):	0·0		0·25			0·50			0·75		Generation
Heritability of indicator trait	h_T^2		0·10	0·25	0·50	0·10	0·25	0·50	0·10	0·25	0·50	interval
Co-heritability	$(h_M r_G h_T)$	0·0	0·04	0·06	0·09	0·08	0·13	0·18	0·12	0·19	0·27	(years)
TEST-SELECTION SYSTEM												
Conventional progeny testing (yield)†		104										6·25
Indicator trait (T) alone												
Normal reproduction. Mass selection		0	20	31	45	40	63	90	60	94	135	3·09
Juvenile MOET scheme. Family index	N‡	0	57	73	88	115	146	175	171	219	263	1·83
	H	0	80	96	111	159	191	223	239	287	334	1·83
Combined selection (T and yield)												
Juvenile MOET scheme. Family index	N	170	179	185	193	196	212	231	221	250	284	1·83
	H	225	238	245	254	260	277	299	293	324	362	1·83
Adult MOET scheme. Family index	N	118	121	124	127	126	131	139	131	141	155	3·67
	H	158	162	165	169	168	174	183	174	185	201	3·67
Efficient progeny testing§												
(1) With initial selection on T of young bulls within full-sibships	N	133	134	135	137	136	138	141	137	141	144	5·58
	H	133	135	136	138	137	140	144	140	144	149	5·58
(2) As (1) with females selected as in a juvenile MOET scheme	N	152	155	157	160	160	165	171	166	174	184	4·17
	H	165	169	172	175	175	181	189	184	194	206	4·17
(3) As (1) with females selected as in an adult MOET scheme	N	139	141	142	144	143	146	150	146	150	157	5·08
	H	154	156	158	161	160	164	169	163	169	177	5·08

† As practised (Nicholas and Smith, 1983).

‡ N — normal ET four males and four females per full sibship, eight donors per sire.
H — high ET eight males and eight females per full sibship, 16 donors per sire.

§ See Table 1.

repeatability of the measurement. The accuracy of selection using information on an indicator trait and on yield can be derived by selection index theory. The procedure is the same as that used by Nicholas and Smith (1983) except that additional terms are included so that the phenotypic and genetic relationships amongst all the available records on both T and M are specified. A detailed example is given in APPENDIX 1 for a juvenile MOET female selection index. Specification of the information included in different indices used in the evaluation of the genetic responses in Table 2 is given in APPENDIX 3. In evaluating the selection responses the generation intervals have been minimized, since this usually gives the highest rates of response, and the intervals used are also shown in Table 2.

The selection indices for females include all the individual and family information on both the indicator trait and on yield, and selection is among individuals on their index rank. For males this might result in selection of more than one male per full-sibship, and so to increased inbreeding. In practice, the restriction to one male per full-sibship could be imposed but it is then difficult to calculate the predicted genetic response. So for males a two stage selection procedure was used first selecting on an index for the best full-sibships, and then selecting within full-sibships on the indicator trait for that individual. Since the within-family information is uncorrelated with the family mean, the genetic responses from the two selection stages can be added. The annual genetic response in phenotypic s.d. units is then

$$\frac{[i_m r_{AIm} + i_w r_{AIw} + i_f r_{AIf}]h_M}{(L_m + L_f)}$$

where i is the standardized selection differential; subscripts m, w and f refer to male full-sibship, within male full-sibship and females respectively; r_{AI} is the correlation between the selected individual's breeding value and the index (I) used in selection; and

L is the generation interval. APPENDIX 2 shows values of the correlations (r_{AI}), for different values of heritability of the indicator trait (T) and for its genetic correlation (r_G) with yield, when derived by the methods described above. The correlation r_{AIw} is given by

$$0{\cdot}5h_T r_G[(n-1)/(n(1-0{\cdot}5h_T^2)]^{1/2}$$

where n is the number of males per full-sib family.

RESULTS

Rates of genetic change

Selection on an indicator trait alone (ignoring yield) allows early selection in both sexes, and is effectively a performance test with mass selection. Despite the shorter generation interval, mass selection on the indicator trait alone only approaches the responses in an efficient progeny test when the co-heritability exceeds the heritability of milk yield (Table 2). Selection for the indicator trait alone using a family index within a juvenile MOET nucleus scheme is more effective and the response in milk yield exceeds that from progeny testing where the co-heritability is only half the heritability of yield.

The most efficient indices in juvenile MOET nucleus schemes will combine information on both yield and the indicator trait. The index that only includes milk yield is more efficient than progeny testing nevertheless further useful gains are obtained with an indicator trait even with low or intermediate levels of co-heritability. This is because the accuracy of selection for milk yield in a juvenile MOET nucleus scheme is not high for milk yield and so there is much scope for improvement by adding further information, thus increasing the response appreciably. The within full-sib male selection makes a substantial contribution to the improvement: taking as an example the results from APPENDIX 2, it provides 0·60 of the extra response when $h_T^2 = 0{\cdot}1$ and $r_G = 0{\cdot}25$ decreasing to 0·38 when $h_T^2 = 0{\cdot}5$ and $r_G = 0{\cdot}75$. For an indicator trait with co-heritability of 0·27 (about the same as the heritability of milk yield) response rates of 0·28 to 0·36 s.d. units per year are possible, corresponding (CV = 0·15) to 4·2 to 5·4% of the mean per year. Similar high results were derived by Christensen and Liboriussen (1986) and these appear to be the highest yet deemed possible in improvement of milk yield.

The increases in response from an indicator trait in combined selection with adult MOET nucleus schemes are more modest. This is because the accuracy of selection for yield in an adult MOET nucleus scheme is already reasonably high (0·55 to 0·65) and further increases in accuracy with an indicator trait are not large. The proportion of the extra response that is derived from within male full-sibship selection is greater in the adult than in the juvenile MOET nucleus schemes. With a co-heritability of 0·27, rates of 0·16 to 0·20 s.d. units per year (2·3 to 3·0% of the mean) are possible.

The rationale for research into finding indicator traits has usually been to give a criterion for an initial selection of young bulls to be progeny tested. Such selection is useful and gives moderate increases in the rates of response. As the initial selection becomes more accurate and effective, the genetic variance in milk yield among the bulls which are progeny tested is reduced. So the response expected (see APPENDIX 4) is not increased as much as might have been thought. The same is true for two-stage progeny testing system in bulls with juvenile and adult MOET breeding in females. Another approach, to use indicator traits to reduce the number of bulls that must be progeny tested and still maintain a high genetic response, has been evaluated by Gill, Pirchner and Schwab (1986).

Blood urea nitrogen (BUN)

Stocks with high and low genetic merit for milk yield provide useful material in the search for indicator traits which may be of use in selection (Hill, 1985). In a series of experiments with calves from lines with high and low genetic merit for milk yield, the levels of BUN after fasting, were significantly higher in the low-yield lines than in the high-yield lines. The details are summarized in Table 3. In experiments 1 and 2, fasting was for a period of 48 h while in experiments 3

and 4 straw was offered *ad libitum* for 5 days. These show co-heritabilities in the range of −0·13 to −0·47. The two trials of Sejrsen, Larsen and Andersen (1984) were with the same calves, and this complicates the pooling of the results. Their disparate results may indicate an effect of age or previous experience of the test, or be due to sampling. APPENDIX 5 derives confidence intervals for the estimates when conditional on the estimated breeding values of the sires used, conservatively assuming the genetic correlation (r_G) is zero. The two experiments with British Friesians have 95% confidence intervals of −0·07 to −0·53, whilst those for the Red Danish are up to half as large again, as a result of having only one sire each per high and low line. Only the results from experiment 4 (Table 3) has 0 within the confidence interval. Pooling results gives an estimate of co-heritability of −0·27(s.e. 0·05). This result should be considered as preliminary. BUN was one of several measurements made in these experiments, and has been selected because of its consistent effect in the four trials.

With the high and low trials it is not possible to separate the genetic correlation from the heritability of the indicator trait, and as seen earlier the extra responses depend on both parameters rather than their product alone. The estimates of co-heritability are those for a single blood sample and could be improved somewhat by multiple sampling, since the repeatability of a BUN sample has been estimated as 0·63 (Tilakaratne, Alliston, Carr, Land and Osmond, 1980). The BUN results look encouraging and if substantiated or even moderated, the use of BUN could add appreciably to rates of genetic improvement in milk yield, as indicated in Table 2.

DISCUSSION

With a given reproductive rate, maximum selection response for a trait is usually a balance between selection accuracy and generation interval. Improvements in any of these will lead to increases in the possible rates of response. With many systems possible, the response for each has to be calculated and the maximum found empirically. Results depend on the specification used and on the restrictions applied by practical application, so the findings are seen to be indicative rather than

TABLE 3

Estimates of co-heritability for milk yield and blood urea nitrogen after fasting, from four experiments on calves with high and low genetic merit for milk yield

		Experiment			
		1	2	3	4
Milk yield (M)					
Phenotypic s.d. (kg)	s_{M2}	(800)†	(800)	(800)	(800)
Heritability	h_M	(0·25)	(0·25)	(0·25)	(0·25)
Estimated difference in genetic merit for yield in calves: high−low line (kg)	R	693	806	(1060)	(1060)
Blood urea nitrogen (T)					
Phenotypic s.d. (mmol/l)	s_T	1·04	0·71	(0·9)	(0·9)
High−low line calves (mmol/l)	CR	−0·92	−0·73	−2·25	−0·60
No. of measurements per calf	n	7	3	3	3
Co-heritability‡	$h_M r_G h_T$	−0·26	−0·26	−0·47	−0·13

† Values in parentheses are assumed.

‡ Falconer (1981) (19·5a) $h_M r_G h_T = \dfrac{CR\, h_M^2 s_M}{R \quad s_T}$

Experiments

1. Tilakaratne *et al.* (1980): four high sires, 21 calves; four low sires, 21 calves; both sexes, 3 to 4 months.
2. Sinnett-Smith *et al.* (1987): four high sires, 17 calves; 10 low sires, 15 calves; both sexes, 4 to 5 months.
3, 4. Sejrsen *et al.* (1984): one high sire, 10 calves, one low sire, six calves; bull calves, 3 to 5 months and 7 months.

precise or absolute. This is particularly so in comparison of MOET nucleus schemes and progeny testing schemes since assumptions of large population sizes (used in estimating selection intensity) are more likely to be met in the latter than the former. Thus, in practice, progress from MOET nucleus schemes may be 10% slower than estimated through size restrictions alone. The results presented are from considering only one indicator trait, but are easily generalized according to standard index selection theory for the case where more than one is available.

Indicator traits can improve response (i) by increasing selection accuracy of indices, (ii) by offering new opportunities for selection where previously random choices were made, and (iii) by reducing generation intervals. It is clear from the results that their benefits in dairy cattle breeding schemes chiefly arise from the first two possibilities and in cases involving MOET a substantial part of the benefit is from the second. This is because the indicator trait allows the discrimination of juvenile full-sib males where other information is identical. The additional accuracy given by the indicator trait may be expected to result in a decrease in the utilizable genetic variance (Bulmer, 1971; Fimland, 1979) but the degree to which this will occur and its consequences, is complex.

In most of the theoretical work with indicator traits (e.g. Sales and Hill, 1976a and b; Hill, 1985) the genetic responses are shown to be functions of the co-heritability alone, and not of the component parts (r_G and h_T). However, in the family indices the equations for full-sibs and half-sibs for the indicator trait have terms including the heritability of the indicator trait (h_T^2), affecting selection accuracy of the individual for the indicator traits, but have no term for the genetic correlation. Thus, h_T and r_G are separated in some of the equations, and though the response depends mainly on their product, there are different responses with different component values. This is explained by recognizing that h_T^2 has been used to provide estimates of the intra-class correlations of half- and full-sibs (this is reasonable for half-sibs but less so for full-sibs) used by Hill (1985). Thus in evaluating potential indicator traits there is no need to design experiments to estimate h_T^2 *per se*, but there is a requirement to estimate the intra-class correlations of sibs.

The theory and calculations assume that the genetic parameters are accurate. Sales and Hill (1976a and b) studied the effects of having inaccurate parameters, and found that the extra predicted response is likely to be reduced by any inaccuracy. They found in the simple case for one trait and one indicator trait, that if the indicator trait was in fact useless as a predictor, the predicted gain in response would in fact be equal to the loss in response incurred by using the useless information. So it would be wise to estimate the effect of an indicator trait well before using it in combined selection, and especially before using it alone in selection. As Sales and Hill (1976a and b) suggest, it is better to be going the right way slowly than the wrong way quickly. However, in the current circumstances the principle of losing response from using useless ($r_G = 0$) information applies only to that part of the predicted gain arising from additional accuracy in selection when combined in indices with milk yield. With MOET, a substantial part of the predicted gain comes from new selection opportunities among full-sibs with otherwise identical information, and if the indicator trait were useless then any choice would be at random for M as it was without the indicator trait.

Of more concern in the use of indicator traits might be the nature of any correlated responses, and the duration of the indirect responses through the indicator trait. Concentrating selection on individual components of production may be effective only temporarily, or may affect other components in a balanced system. Concern over unfavourable correlation responses has prevented the development of schemes using rates of thyroxine degradation in young bulls as predictors of genetic merit for yield (Joakimsen, Steenberg, Lien and Theodorsen, 1971; Sørensen, Kruse and Andersen, 1981). Although both studies reported a significant genetic correlation with milk yield of 0·42, concern over possible effects on metabolic

rate, and so perhaps on the efficiency of production, has led to detailed long-term trials being carried out before developing thyroxine degradation rate as a predictor in practice. This seems a very conservative attitude, and not one to be recommended with BUN. A more positive approach should be adopted; first to check on the co-heritability in independent material, secondly to test it in a section of the main breeding population to measure its value in practice, and thirdly to develop small high and low lines for the indicator trait to study direct and correlated responses.

Juvenile MOET nucleus schemes were indicated as having not only the maximum rates of progress without using indicator traits but also the greatest potential for their use. However, these schemes also tend to have higher rates of inbreeding than progeny testing schemes or adult MOET nucleus schemes with the same facilities. This arises through the greater reliance on family indices (Hill, 1972; Burrows, 1984) and the small population sizes of such nucleus herds due to the cost of MOET. Although increasing accuracy the use of indicator traits in both adult and juvenile MOET nucleus schemes will ease the problem associated with the use of family indices, since more weight is given to information from the candidate (i.e. female or male full-sibship) itself rather than from its relatives. The degree of easing will depend on the genetic parameters. Regardless of whether indicator traits are in use, the inbreeding problem would be reduced if several schemes were run concurrently and exchanged breeding stock, a solution that the potential benefits in national terms would warrant. A modification of the MOET schemes that would also help is to use more than one or all males per sibship, but at proportionately lower mating ratio. However, this would reduce the accuracy of selection slightly by having smaller half-sib families, and, more importantly, in increasing the proportion of males selected it reduces the opportunity for utilizing the within full-sib family genetic variation by indicator traits. A further problem caused by using family indices on small populations is the reduction of selection differentials or genetic variance compared with those in large populations (Rawlings, 1976; Hill, 1976 and 1977). Recent simulations in a small adult MOET scheme (Juga and Maki-Tanila, 1987) have found lower than expected selection differentials through the domination of a small number of large families. As with inbreeding rate, this problem would be eased by the use of indicator traits and by the use of more than one male per sibship to increase the number of families but, of course, the latter option has the same drawbacks as before.

In conclusion, the important result is the high rate of genetic change possible with juvenile MOET nucleus schemes, and the enhancement that indicator traits can make to a variety of breeding schemes, but to juvenile MOET nucleus schemes in particular. Moreover, these theoretical results may be ripe for development in practice as potentially valuable indicator traits, such as BUN, are now being identified. The results demonstrate the challenge for research workers in different fields from reproductive physiology, to metabolic physiology, to realize the potential.

REFERENCES

AVALOS, E. and SMITH, C. 1987. Genetic improvement of litter size in pigs. *Animal Production* **44:** 153-163.

BRASCAMP, E. W. 1978. Methods on economic optimisation of animal breeding plans. *Research Institute of Animal Husbandry, Zeist, Netherlands*, pp. 70-72.

BULMER, M. G. 1971. The effect of selection on genetic variability. *American Naturalist* **105:** 201-211.

BURROWS, P. M. 1984. Inbreeding under selection from related families. *Biometrics* **40:** 895-906.

CHRISTENSEN, L. G. 1984. Breeding problems in small cattle populations. *35th Meeting of the European Association of Animal Production, The Hague, Netherlands.*

CHRISTENSEN, L. G. and LIBORIUSSEN, T. 1986. Embryo transfer in the genetic improvement of dairy cattle. In *Exploiting New Technologies in Animal Breeding: Genetic Developments* (ed. C. Smith, J. W. B. King and J. C. McKay), pp. 37-46. Clarendon Press, Oxford.

COLLEAU, J. J. 1985. [Genetic improvement by embryo transfer within selection nuclei in dairy cattle.] *Génétique Sélection et Evolution* **17:** 499-537.

FALCONER, D. S. 1981. *Introduction to Quantitative Genetics*. 2nd ed. Longman, London.

FIMLAND, E. 1979. The effect of selection on additive genetic parameters. *Zeitschrift fur Tierzuchtung und Zuchtungsbiologie* **96:** 120-134.

GILL, J. L., PIRCHNER, F. and SCHWAB, M. 1986. Discriminating breeding values of young bulls from their serum concentrations of hormones. *Journal of Animal Breeding and Genetics* **103:** 241-248.

HILL, W. G. 1972. Estimation of genetic change. 1. General theory and design of control populations. *Animal Breeding Abstracts* **40:** 1-15.

HILL, W. G. 1976. Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32:** 889-902.

HILL, W. G. 1977. Order statistics of correlated variables and implications in genetic selection programmes. II. Response to selection. *Biometrics* **33:** 703-712.

HILL, W. G. 1985. Detection and genetic assessment of physiological criteria of merit within breeds. In *Genetics of Reproduction in Sheep* (ed: R. B. Land and D. W. Robinson), pp. 319-331. Butterworths, London.

JOAKIMSEN, Ø, STEENBERG, K., LIEN, H. and THEODORSEN, L. 1971. Genetic relationship between thyroxine degradation and fat-corrected milk yield in cattle. *Acta Agriculturae Scandinavica* **21:** 121-124.

JUGA, J. and MAKI-TANILA, A. 1987. Genetic change in nucleus breeding daily herds using embryo transfer. *Acts Agriculturae Scandinavica* **37:** 511-519.

LAND, R. B. and HILL, W. G. 1975. The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Animal Production* **21:** 1-12.

NICHOLAS, F. W. 1979. The genetic implication of multiple ovulation and embryo transfer in small dairy herds. *30th Meeting European Association of Animal Production, Harrogate, Paper G1.11.*

NICHOLAS, F. W. and SMITH, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

OWEN, D. B. and STECK, G. P. 1962. Moments of order statistics from the equicorrelated normal distribution. *Annals of Mathematical Statistics* **33:** 1286-1291.

RAWLINGS, J. O. 1976. Order statistics for a special class of unequally correlated multinormal varieties. *Biometrics* **32:** 875-887.

SALES, J. and HILL, W. G. 1976a. Effect of sampling errors on the efficiency of selection indices. 1. Use of information from relatives for single trait improvement. *Animal Production* **22:** 1-17.

SALES, J. and HILL, W. G. 1976b. Effect of sampling errors on the efficiency of selection indices. 2. Use of information on associated traits for improvement of a single important trait. *Animal Production* **23:** 1-14.

SEJRSEN, K., LARSEN, F. and ANDERSEN, B. B. 1984. Use of plasma hormone and metabolic levels to predict breeding value of young bulls for butterfat production. *Animal Production* **39:** 335-344.

SINNETT-SMITH, P. A., SLEE, J. and WOOLLIAMS, J. A. 1987, Biochemical and physiological responses to metabolic stimuli in Friesian calves of differing genetic merit for milk production. *Animal Production* **44:** 11-19.

SØRENSEN, M. T., KRUSE, V. and ANDERSEN, B. B. 1981. Thyroxine degradation rate in young bulls of Danish dual-purpose breeds. Genetic relationship to weight gain, feed conversion and breeding value for butterfat production. *Livestock Production Science* **8:** 399-406.

TILAKARATNE, N., ALLISTON, J. C., CARR, W. R., LAND, R. B. and OSMOND, T. J. 1980. Physiological attributes as possible selection criteria for milk production. 1. Study of metabolites in Friesian calves of high and low genetic merit. *Animal Production* **30:** 327-340.

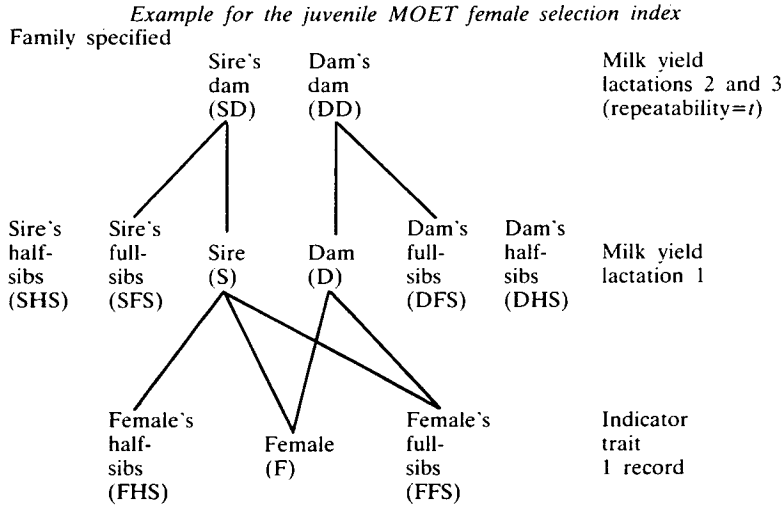
WALKLEY, J. R. W. and SMITH, C. 1980. The use of physiological traits in genetic selection for litter size in sheep. *Journal of Reproduction and Fertility* **59:** 83-88.

WIENER, G., SINNETT-SMITH, P. A., SLEE, J. and WOOLLIAMS, J. A. 1986. Genetic variation in physiological factors relating to livestock improvement. *Report of AFRC Animal Breeding Research Organisation*, pp. 14-16.

(*Received* 24 *February* 1987—*Accepted* 24 *November* 1987)

APPENDIX 1

Example for the juvenile MOET female selection index



Family size: n males and n females per full-sibship; s full-sibships per sire.

Parameter: Heritabilities; milk yield $h_M^2 = M$, indicator trait $h_T^2 = T$.
Genetic correlation between milk yield and indicator trait, r_G.
Co-heritability $h_M r_G h_T = C$.
Intra-class correlation estimates; milk yield, half-sibs $M/4$, full-sibs $M/2$; indicator trait, half-sibs $T/4$, full-sibs $T/2$.

Phenotypic matrix (P) *when* M *and* T *are standardized so that* $s_M^2 = s_T^2 = 1$

		1	2	3	4	5	6	7	8	9	10
F	1	1	$0{\cdot}5T$	$0{\cdot}25T$	$0{\cdot}5C$	$0{\cdot}25C$	$0{\cdot}125C$	$0{\cdot}25C$	$0{\cdot}125C$	$0{\cdot}25C$	$0{\cdot}25C$
FFS	2		P_2	$0{\cdot}25T$	$0{\cdot}5C$	$0{\cdot}25C$	$0{\cdot}125C$	$0{\cdot}25C$	$0{\cdot}125C$	$0{\cdot}25C$	$0{\cdot}25C$
FHS	3			P_3	0	0	0	$0{\cdot}25C$	$0{\cdot}125C$	0	$0{\cdot}25C$
D	4				1	$0{\cdot}5M$	$0{\cdot}25M$	0	0	$0{\cdot}5M$	0
DFS	5					P_5	$0{\cdot}25M$	0	0	$0{\cdot}5M$	0
DHS	6						P_6	0	0	0	0
SFS	7							P_7	$0{\cdot}25M$	0	$0{\cdot}5M$
SHS	8								P_8	0	0
DD	9									P_9	0
SD	10										P_{10}

$P_2 = (1+(2n-2)0{\cdot}5T)/(2n-1)$
$P_3 = (1+(2n-1)0{\cdot}5T + 2n(s-2)0{\cdot}25T)/[2n(s-1)]$
$P_5 = (1+(n-2)0{\cdot}5M)/(n-1)$
$P_6 = (1+(n-1)0{\cdot}5M + n(s-2)0{\cdot}25M)/[n(s-1)] = P_8$
$P_7 = (1+(n-1)0{\cdot}5M)/n$
$P_9 = (1+t)/2 = P_{10}$

Genetic vector (**g**) standardized similarly to **P**
$(C, 0{\cdot}5C, 0{\cdot}25C, 0{\cdot}5M, 0{\cdot}25M, 0{\cdot}125M, 0{\cdot}25M, 0{\cdot}125M, 0{\cdot}25M, 0{\cdot}25M)$.

Selection index weights **b** are given by $\mathbf{b} = \mathbf{P}^{-1}\mathbf{g}$.

The correlation of the selection index with the breeding value for milk yield of the female is

$$\mathbf{b'g}\ (\mathbf{b'Pb}\ h_M^2)^{-1/2}$$

Example
$h_M^2 = M = 0{\cdot}25$
$h_T^2 = T = 0{\cdot}50$
$r_G = 0{\cdot}50$
$C = 0{\cdot}50 \times (0{\cdot}25 \times 0{\cdot}5)^{1/2} = 0{\cdot}1768$
$n = 4$
$s = 8$
Correlation = 0·526.

APPENDIX 2

Example of accuracy and response (R) *in selection for milk yield by selection indices including an indicator trait, for a juvenile MOET scheme (normal ET rates, four males and four females per sibship, eight donors per sire, generation interval 1·83 years)*

Heritability of the indicator trait (h^2_I)		Genetic correlation (r_G) of indicator trait and yield			
		0·00	0·25	0·50	0·75
0·10	r_{AIm}†	0·427	0·436	0·463	0·508
	r_{AIw}	0·000	0·035	0·070	0·105
	r_{AIf}‡	0·427	0·437	0·468	0·519
	R	0·170	0·179	0·196	0·221
0·25§	r_{AIm}		0·440	0·480	0·547
	r_{AIw}		0·058	0·116	0·174
	r_{AIf}		0·444	0·494	0·573
	R		0·185	0·212	0·250
0·50	r_{AIm}		0·444	0·495	0·581
	r_{AIw}		0·088	0·177	0·265
	r_{AIf}		0·453	0·526	0·639
	R		0·193	0·231	0·284

† Correlation of breeding value (G) for milk with:

		Proportion selected	Selection differential
Full-sib family selection index for males	r_{AIm}	1/8	1·647
Indicator trait *within* full-sib family for males	r_{AIw}	1 from 4	1·188
Full selection index for females (see APPENDIX 1)	r_{AIf}	1/4	1·271

‡ Estimated genetic response (R)

$$\frac{(1{\cdot}647\ r_{AIm} + 1{\cdot}188\ r_{AIw} + 1{\cdot}271\ r_{AIf})h_M \text{ s.d. units.}}{(1{\cdot}83 + 1{\cdot}83)}$$

§ The values for $r_G = 0$ are constant for all values of h^2_I.

APPENDIX 3

Details of information used in the selection systems evaluated in Table 2

	Selection system									
	Indicator trait alone		Combined selection							
							Progeny testing for the best male for the indicator trait within each full-sibship			
							Females selected as in			
	Normal reproduction	Juvenile MOET	Juvenile MOET		Adult MOET		Juvenile MOET		Adult MOET	
	Mass selection	Family index	m†	f	m	f	m	f	m	f
Indicator trait										
Individual	+	+	+	+	+	+	+	+	+	+
Full-sibs		+	+	+	+	+	+	+	+	+
Half-sibs		+	+	+	+	+		+		+
Yield										
Individual						+				+
Full-sibs					+	+				+
Half-sibs					+	+				+
Dam			+	+	+	+		+		+
Dam's full-sibs			+	+				+		
Dam's half-sibs			+	+				+		
Dam's dam			+	+				+		
Sire's full-sibs			+	+				‡		
Sire's half-sibs			+	+				‡		
Sire's dam			+	+				‡		
Progeny							+		+	

† m = male; f = female.

‡ Accurate selection of sires on progeny test, so little genetic variation left in this path.

APPENDIX 4

Two stage selection of males for progeny testing

1. Initial selection; on indicator trait within male full-sibships of size n:
Genetic difference $= i_1\ r_w h_M s_M$
where $r_w = 0{\cdot}5\ h_T r_G \sqrt{\dfrac{(n-1)}{n(1-0{\cdot}5h_T^2)}}$
and the selection intensity
$i_1 = 1{\cdot}188$ for $n=4$ and $1{\cdot}522$ for $n=8$.

2. Subsequent selection; on progeny test of m offspring:
Genetic difference $= i_2 h_M s_M \sqrt{\beta[1-(1-k_n)r_w^2]}$
where $\beta = \dfrac{0{\cdot}25(1-(1-k_n)r_w^2)h_M^2}{[0{\cdot}25(1-(1-k_n)r_w^2)h_M^2 + (1-0{\cdot}25h_M^2)/m]}$

and k_n represents the variance of the highest value of a sample of size n from a $N(0,1)$ distribution with pairwise correlations of $-1/(n-1)$. The variance of the greatest deviation from the means of a sample of n full-sibs is thus $k_n \dfrac{(n-1)}{n}(1-0{\cdot}5h_T^2)$. For $n=4$, $k_n=0{\cdot}32$ and $n=8$, $k_n=0{\cdot}29$.

The value of i_2 used assumes that the distribution for the second stage of selection can still be approximated by a normal distribution despite alteration through the initial selection. The total genetic difference is the sum of the genetic differences at the two stages.

APPENDIX 5

Confidence limits for the co-heritability estimates

Assume without loss of generality that the traits (milk yield and BUN) are standardized so that $s_M^2 = s_T^2 = 1$, the heritabilities are h_M^2 and h_T^2 respectively, the genetic correlation is r_G and the repeatability of BUN is t. Suppose further the estimated Improved Contemporary Comparisons (ICC) ($0{\cdot}5 \times$ breeding value of sires are known ($\hat{S}_k$ for sire k) and were estimated by regressing its progeny mean by β where $\beta = [0{\cdot}25h_M^2]/[0{\cdot}25h_M^2+(1-0{\cdot}25h_M^2)/m]$ where m is the effective number of daughters. Then for two distinct half-sibs i, j offspring of sire k, with G_i, G_j their genetic merit for milk yield and T_i, T_j the phenotypes for BUN after n measurements.

$$\begin{aligned}
\mathrm{Var}(G_i|\hat{S}_k) &= 0{\cdot}75h_M^2 + 0{\cdot}25h_M^2(1-\beta)\\
\mathrm{cov}(G_i,G_j|\hat{S}_k) &= 0{\cdot}25h_M^2(1-\beta)\\
\mathrm{Var}(T_i|\hat{S}_k) &= ([1+(n-1)t]/n - 0{\cdot}25h_T^2) + 0{\cdot}25h_T^2(1-r_G^2\beta)\\
\mathrm{cov}(T_i,T_j|\hat{S}_k) &= 0{\cdot}25h_T^2(1-r_G^2\beta)\\
\mathrm{cov}(G_i,T_i|\hat{S}_k) &= 0{\cdot}75h_M r_G h_T + 0{\cdot}25h_M r_G h_T(1-\beta)\\
\mathrm{cov}(G_i,T_j|\hat{S}_k) &= 0{\cdot}25h_M r_G h_T(1-\beta)
\end{aligned}$$

Then for the experiment of Tilakaratne *et al.* (1980)

$$\begin{aligned}
\mathrm{Var}(G_H-G_L|\hat{S}_H,\hat{S}_L) &= 0{\cdot}071h_M^2+0{\cdot}126h_M^2(1-\beta)\\
\mathrm{Var}(T_H-T_L|\hat{S}_H,\hat{S}_L) &= 0{\cdot}095([1+(n-1)t]/n - 0{\cdot}25h_T^2)+0{\cdot}126h_T^2(1-r_G^2\beta)\\
\mathrm{cov}(G_H-G_L,T_H-T_L|\hat{S}_H,\hat{S}_L) &= 0{\cdot}071h_M r_G h_T + 0{\cdot}126h_M r_G h_T(1-\beta)
\end{aligned}$$

where G_H,G_L, T_H and T_L are the genetic and phenotypic means of calves of high and low merit and S_H and S_L are the order statistics of the estimated ICC's of high and low sires. From these an approximate confidence interval can be obtained for the estimate of $h_M r_G h_T = h_M^2 \dfrac{(T_H-T_L)}{(G_H-G_L)}$ from Fieller's Theorem (Fieller, 1944) using the estimate of $t = 0{\cdot}63$ (Tilakaratne *et al.*, 1980) and with assumptions $\beta = 0{\cdot}75$ (i.e. $m = 45$) and $h_T^2 = 0{\cdot}4$.

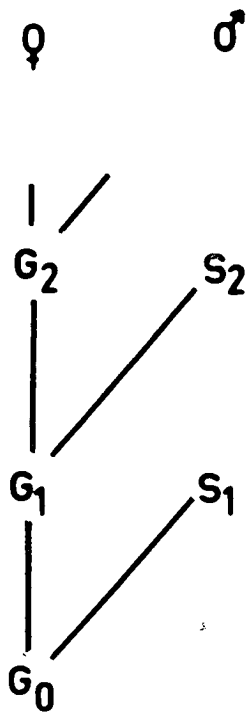


FIG. 1. A diagram showing pedigree for a calf in the experiment of Sinnett-Smith *et al.* (1987).

The experiment of Sejrsen *et al.* (1984) follows similarly.

For the experiment of Sinnett-Smith *et al.* (1987) this approach is modified since calves were produced from repeated selection for high and low yield. Bulls used were from a panel of low, or of high ICC. Heifers used as replacements entered the line determined by their sire. Estimated genetic merit of the calf is derived solely from the estimated ICC's of male ancestors. Consider the pedigree shown in Figure 1. Since inbreeding is avoided it is assumed that bulls used are distinct.

$$\begin{aligned}\mathrm{Var}(G_0|\hat{S}_1,S_2,\hat{S}_3\ldots\hat{S}n) &= 0{\cdot}5h^2_M+0{\cdot}25h^2_M(1-\beta)+0{\cdot}25\mathrm{Var}(G_1|\hat{S}_1,\hat{S}_2,\hat{S}_3\ldots\hat{S}n)\\ &= 0{\cdot}5h^2_M+0{\cdot}25h^2_M(1-\beta)\\ &\quad+0{\cdot}25[0{\cdot}5h^2_M+0{\cdot}25h^2_M(1-\beta)+0{\cdot}25\mathrm{Var}(G_2|\hat{S}_2,\hat{S}_3\ldots\hat{S}n)]\end{aligned}$$

$$\begin{aligned}\text{Thus } \mathrm{Var}(G_0|\hat{S}_1,\hat{S}_2,\hat{S}_3\ldots\hat{S}n) &\to [0{\cdot}5h^2_M+0{\cdot}25h^2_M(1-\beta)]\sum_{i=0}4^{-i}\\ &= h^2_M(3-\beta)/3 \text{ as } n\to\infty\end{aligned}$$

Other results then follow in a similar manner and, ignoring common ancestry other than sires, confidence intervals are derived as before.

ADDITIONAL REFERENCES

FIELLER, E. C. 1944. A fundamental formula in the statistics of biological assay and some applications. *Quarterly Journal of Pharmacy and Pharmacology* **17**: 117-123.

Paper 2

STRATEGIES TO MAXIMISE SELECTION PROGRESS IN DAIRY CATTLE

J.A. Woolliams
AFRC Institute of Animal Physiology and Genetics Research, Roslin, Midlothian EH25 9PS

SUMMARY

The basic principles for maximising progress using the infinite alleles model are largely understood. However this is not the case when individual alleles can be identified in conjunction with continuous variation, nor when constraints on inbreeding are applied. Accounting for linkage disequilibrium does not alter the relative merits of progeny testing and MOET nucleus schemes. The reduction in the rate of progress at equilibrium in progeny testing is similar to the reduction in the total genetic variance. However in MOET nucleus schemes the reduction in progress is greater than the reduction in genetic variance alone since the value of ancestral information is also reduced by linkage equilibrium.

BLUP selection enhances progress and removes the distinctions between MOET nucleus schemes and progeny testing schemes. . When the female family size is increased, factorial mating gives extra progress when constrained by inbreeding. When accuracy is low, as in MOET nucleus schemes or the use of young unproven bulls indicator traits would prove useful. Cloning has potential for reducing genetic lags, but not as yet for increasing the progress of leading herds; sexing of embryos and semen can improve progress through more effective use of the allocated resources.

INTRODUCTION

The maximisation of genetic progress has become more and more tractable for animal breeders over the last decade. This has principally been achieved through the development of statistical and computational techniques for estimating breeding values. However the single-minded pursuit of progress is not without its risks and consequences - not only in management, but through the effects of more rapid inbreeding and loss of genetic variation for production and disease resistance.

Therefore the first section of the paper reviews the basic principles upon which genetic progress is maximised, together with the risks and consequences and the second section will examine how genetic progress may be maximised further, risks notwithstanding, given the current procedures for evaluation and the likely developments over the next decade.

BASIC PRINCIPLES

The selection objective may be complex and composed of a range of traits of varying economic importance but it will be assumed that it is clearly defined. This problem has historically led to 3 selection procedures to achieve the objective; tandem selection whereby each trait is selected on in turn; multiple culling levels whereby breeding livestock have to satisfy threshold requirements on each trait; or, index selection in which measurements of each trait are combined linearly to give an aggregate value upon which selection is then practised. It has been clearly shown that for linear economic functions of the traits the efficiency of these procedures rank 'index', 'multiple culling' and 'tandem' in descending order. Thus it is sufficient in discussing the maximum selection progress to consider progress in a single trait i.e. the aggregate genotype.

Progress assuming the infinite-loci model

Throughout the phenotypes of the base population (t=0) are assumed to be normally distributed with variance $V_P(0)=1$ and unless otherwise stated the infinite-loci model of Bulmer (1980) with no linkage or epistasis will be assumed. The breeding values of the base

population are assumed normally distributed with variance $V_A(0)=h^2$ and under the model the regression of the breeding value of an individual (A_c) measured prior to selection on its parents (A_s and A_d) remains invariant with

$$A_c = 1/2\ A_s + 1/2\ A_d + A_m \qquad (1)$$

where A_m is a Mendelian sampling term normally distributed with mean zero and variance $1/2\ h^2\ (1\text{-}F)$ and F is the mean inbreeding coefficient of its parents. The regression of breeding value on phenotype (Y) in generation t is

$$A = h^2(t)\ Y + e \qquad (2)$$

where $h^2(t) = V_A(t)/V_P(t)$ and e is normally distributed with mean zero and variance $(1\text{-}h^2(t))V_A(t)$. Under normality this regression is preserved by truncation of the population.

Genetic progress in generation t is given by

$$\Delta G(t) = (i_M\ r_{IAM}(t) + i_F\ r_{IAF}(t))\sqrt{V_A(t)}/(L_M + L_F) \qquad (3)$$

where i denotes standardised selection differential, r_{IA} selection accuracy (the correlation between the index used and the true aggregate breeding value), L denotes generation interval and the subscripts M and F denote males and females respectively. Strictly speaking, for overlapping generations this defines the asymptotic rate of progress but this will suffice for the current purpose.

Selection progress is thus increased by increasing selection differential, i.e. by decreasing selection proportion to a limit that is determined by the practical reproductive rate of the male and female. It is principally by this route that the major improvements in selection progress were made over the last 40 years through the increase in male reproductive rate made possible by AI and the freezing of semen, and where more modest gains are possible through the use of ET to increase the female reproductive rate. Secondly the accuracy of estimating the breeding value can be increased through i) controlling the environment, ii) reducing measurement errors, iii) good design of testing schemes, iv) incorporating as much information on relevant traits and on relatives as possible, and v) by fitting as good models as possible. The use of information from relatives and computational power to fit appropriately large statistical models are ways in which BLUP and individual animal models will help speed genetic progress. However increasing gain through accuracy cannot be viewed in isolation: for example, accuracy on females can be improved by recording many lactations before breeding decisions are made but this in turn increases the generation interval and hence from (3) reduces progress. Thus there is a compromise to be made between accuracy and generation interval, and this lies at the heart of much of the debate on ways to maximise genetic gain in dairy cattle (e.g. MOET and progeny testing).

Finally genetic progress can be increased by increasing V_A. At first sight this may seem impossible given the base population and also given the selected parents. However by the use of positive assortative mating genetic variation in the offspring can be increased. From (1) for any generation

$$V_A(t) = 1/4\ (V_{AS} + V_{AD}) + 2\ \mathrm{cov}\ (A_S, A_D) + 1/2\ h^2\ (1\text{-}F)$$

where V_{AS} and V_{AD} are the variance of breeding values amongst selected sires and dams of the previous generation. Thus by introducing a positive covariance between the expected

breeding values of the mates and thus, in consequence, between the true breeding values of the mates, V_A is increased. By analogy it is also true that if the covariance is negative, through negative assortative mating, then progress will be reduced.

Smith and Hammond (1987) confirmed the work of others that assortative mating is particularly useful when h^2 is high and selection intensity is low, but extra progress exceeds 10% only when $h^2 \nearrow 1$ and $i \searrow 0$. These authors questioned the value of assortative mating when distributions are non-normal. However it can be shown (proof not given) that positive assortment is never counter-productive when

i) for each sex, prediction error variances of expected breeding values for individuals are identically distributed.
ii) the breeding value of the offspring is the average of its parents with deviation distributed in any form but independent of the parental breeding values.
iii) inbreeding is either negligible or uniform in the population.

Assumption (ii) include (1) as a special case. Thus if there are two 'discordant' pairs (S1,D1) and (S2,D2) i.e. with expected breeding values $E(A_{S1}) < E(A_{S2})$ but $E(A_{D1}) > E(A_{D2})$, then selection progress is at least as good if the pairs are rearranged to be concordant.

Linkage disequilibrium

The infinite model predicts that truncation selection will reduce the genetic variation present in the population (Bulmer, 1980). Let an asterisk denote variance after selection and I(t) the regressed index of selection.

$$V_A(t)^* = V_I(t)^* + (1-r_{IA}^2(t))V_A(t)$$

and from (1), with $k = i(i-x)$ so that $V_I(t)^* = (1-k)r_{IA}^2(t)V_A(t)$

$$V_A(t+1) = 1/2(h^2 + V_A(t)) - 1/4(k_M r_{IAM}^2(t) + k_F r_{IAF}^2(t))V_A(t)$$

Thus $V_A(1) < V_A(O)$ and Bulmer continues to show that the recurrence relation (4) converges rapidly for truncation selection. Thus most genetic variation is predicted to disappear in the first round of selection. From the assumptions of the infinite model the decline in genetic variance can only be explained by linkage disequilibrium induced by selection. Furthermore, the model predicts that upon ceasing selection genetic variance is progressively restored, the shortfall reducing by 50% per generation.

The effects of the linkage disequilibrium in the infinite model are i) loss of progress through lower $V_A(t)$ and ii) a reduction in the correlation between individual performance and that of its sibs due to a smaller contribution of genetic variance from the parents.

Since increasing selection intensity increases progress but also reduces genetic variation and, as a consequence, subsequent genetic progress, there is an a priori case that maximal selection progress may be obtained from a trade-off between i and $V_A(t)$. However, it can be shown (proof not given) that under a model described by (1) with the same conditions given for assortative mating (hence including the infinite model) selection progress is always maximised by selecting the best individuals.

Progress with finite-loci models

It is unfortunate that the clear results for the infinite-loci models do not apply generally to models with finite loci. Examples can be found to show i) positive assortative

mating is not always beneficial; ii) $V_A(t)$ need not decrease with selection; iii) progress is not greatest by selecting the best. The key to all the conditions for the results in the infinite-loci model is the homogenous distribution of the offspring about the mean expected breeding value of their parents, a condition that is broken by the Mendelian sampling terms when specific alleles are considered in homozygous or heterozygous conditions. Thus strategies for maximising gain will depend specifically on gene frequencies and gene effects. The optimal strategies for selection when quantitative variation is combined with even one major gene are also unclear. This situation may become a reality of proposals to map the complete bovine genome are pursued.

Inbreeding

Unlike linkage disequilibrium, whatever the allelic model, inbreeding causes a permanent reduction in the genetic variance of a homogeneous population. It is this loss that ultimately brings genetic progress to a halt, apart from new mutational variance arising. Inbreeding is unavoidable in finite populations and is usually, but not always, exacerbated by genetic selection. All the methods for increasing progress described previously such as higher selection intensities, including the use of ancestral information and assortative mating will increase inbreeding when used with mass selection.

Thus from the narrow viewpoint of future selection progress there is a need to trade off progress with inbreeding. Faster inbreeding will also cause greater drift in neutral traits some of which may become of economic importance as a result (e.g. susceptibility to disease) and will also lead to a greater variance in the response of the selected traits. Furthermore there is the economic loss in the present to consider. Hudson and Van Vleck (1984) although suggesting that current levels of inbreeding were not of concern, estimated a 21kg drop in milk yield per cent inbreeding with reductions in longevity and calving interval. The possible hidden costs of reduced growth and greater disease and mortality, well-documented in other species, were not measured.

The consideration of inbreeding introduces a further concept: the time horizon over which genetic gain is to be maximised - one generation, n generations for some n, or a weighted average over generations such as would be produced by discounted benefit techniques. Smith (1969) considered adjusting intensity of mass selection to maximise response over a fixed n generations and Demple (1974) showed that when $n \nearrow \infty$ so response is greatest using within-family selection rather than mass selection.

PROGRESS IN THE DAIRY POPULATION

Two major themes over the last decade concerning dairy cattle have been i) the relative advantages of MOET nucleus herds and progeny testing and ii) BLUP and individual animal models.

MOET and progeny testing

Progeny testing is a robust system of improvement that overcomes the sex-limitation of dairy improvement. It is characterised by selecting bulls with high accuracy but with a long generation interval. Nicholas and Smith (1983) and Woolliams and Smith (1988) examined what impact multiple ovulation and embryo transfer could have in a nucleus herd. By increasing the reproductive rate of the cow selection intensity could be increased. Moreover the increase in the number of close relatives (principally full-sibs) meant that information on their performance could help in increasing the accuracy of female selection. For males, the extra information from full-sib sisters provided more information on the breeding value of its dam, whilst the female paternal half-sibs provided information on its

sire. Their conclusions were that sufficient information was obtained by 4 years of age, after the lst lactation of the females, to allow selection to take place (thus reducing the generation interval below that of progeny testing) and to give rates of change within the nucleus comparable to those from progeny testing. This was termed the 'Adult' (A) Scheme. Furthermore these authors showed that by selecting progeny at 2 years of age using ancestral information (i.e. aunts, half-aunts and grand-dams) and then using MOET, progress was further enhanced. This was termed the 'Juvenile' (J) scheme, and can be characterised by low accuracy and low generation interval, with the Adult scheme characterised by medium accuracy and medium generation interval. This is an example of the trade-off between r_{AI} and L required to maximise progress.

Since the publication of these results it has been a subject of dispute as to whether or not the advantages given by Woolliams and Smith (1988) could be realised in practice. Simulations carried out (Juga and Maki-Tanila, 1987; and latterly Ruane and Thompson, 1990) suggest that progress obtained in A schemes is perhaps only 60-65% of that proposed.

A major cause of the shortfall was identified as a loss of between family variation i.e. linkage disequilibrium. Proponents of MOET had not explicitly accounted for the loss in genetic variation and, furthermore, MOET uses ancestral and collateral information whose value is also reduced by linkage disequilibrium. However a fair comparison of progeny testing and MOET must also take into account the loss of progress below initial expectation that is bound to occur in models of progeny testing. This has been shown by simulation (Meyer and Smith, 1990) and by complex deterministic models (Meuwissen, 1989). These systems have therefore been compared again using the original structures of Woolliams and Smith (1988), with progress estimated both initially and in equilibrium (which is quickly reached before inbreeding would become a serious consideration). Results are shown in Table 1; two values of h^2 (0.25 and 0.35) are used since the published estimates of progress may be appropriate if estimates of h^2 used are interpreted as coming from a base population already under selection using progeny testing.

Table 1 Rates of progress (base phenotypic s.d./annum) in progeny testing and MOET nucleus schemes

		(1) Initial parameters		(2) Equilibrium parameters		(2)/(1)	
Heritability (h^2)		0.25	0.35	0.25	0.35	0.25	0.35
Progeny testing	A*	0.133	0.166	0.100	0.125	0.75	0.75
	B	0.133	0.166	0.102	0.125	0.77	0.75
Adult MOET	N+	0.118	0.151	0.092	0.115	0.78	0.76
	H	0.158	0.198	0.120	0.148	0.76	0.75
Juvenile MOET	N	0.170	0.218	0.134	0.169	0.79	0.77
	H	0.225	0.282	0.175	0.217	0.78	0.77

*A = 50 effective daughters; B = bulls tested to constant accuracy
+N = 8 calves/donor, 8 donors/sire; H = 16 calves/donor, 16 donors/sire

The relative rates of progress predicted are very similar whichever set of parameters is used, all being reduced by 20-25% when linkage disequilibrium is introduced. Increasing h^2 exacerbates the effect of disequilibrium in all schemes. Table 2 shows $V_A(t)$ is reduced most in progeny testing, and least in juvenile schemes. However whilst progress in progeny testing is reduced in line with the true h^2 at equilibrium ((3) in Table 2), progress in MOET schemes is further reduced, particularly in J schemes. This reflects the reliance of MOET schemes on ancestral or collateral performance rather than individual or offspring performance. The reduction in value of ancestral information can be judged by the extra reduction in h^2 estimated from paternal half-sibs in ((4) Table 2) compared to the true heritabilities at equilibrium. The reduction in progress of J schemes, relying as they do entirely on ancestral information, is very similar to the reduction in the half-sib estimate of h^2. A schemes are intermediate.

Table 2 Genetic variance at equilibrium ($V_A(\infty)$)

		(1)$V_A(\infty)$*		(2)$V_A(\infty)/h^2$		(3)$h^2(\infty)/h^2$		(4)Column(3) from paternal half-sibs	
Heritability (h^2)		0.25	0.35	0.25	0.35	0.25	0.35	0.25	0.35
Progeny testing	A*	0.184	0.250	0.74	0.71	0.79	0.79	0.74	0.72
	B	0.180	0.250	0.72	0.71	0.78	0.79	0.72	0.72
Adult MOET	N*	0.206	0.280	0.82	0.80	0.86	0.86	0.72	0.70
	H	0.199	0.272	0.80	0.78	0.84	0.84	0.67	0.66
Juvenile MOET	N	0.228	0.314	0.91	0.90	0.93	0.93	0.84	0.82
	H	0.224	0.310	0.89	0.88	0.92	0.92	0.81	0.80

*See Table 1; *For cows

The equilibrium half-sib estimate of h^2 in progeny testing when initial $h^2 = 0.35$ is very close to 0.25, suggesting that simulation studies for milk production might be more appropriate using 0.35 as the heritability of an unselected base population rather than 0.25.

In summary, the relative merits of progeny testing and MOET schemes are not greatly affected by linkage disequilibrium. However it is now possible to match the simulations of A schemes to the deterministic predictions: this study suggests a 22% reduction through linkage disequilibrium and Woolliams (1989) found a 15% loss through the finite correlated selection intensities - putting those together suggests that in the absence of inbreeding only 0.66 of the predicted progress will be realised, in close accord with the simulations.

BLUP and individual animal models

For the purposes of this paper BLUP methodology will not be discussed. Some of its key properties are i) the unbiassed predictions of breeding value which means that for any individual, no matter how good the predictions are, further information is as likely to increase the prediction as decrease it, ii) the simultaneous estimation of genetic and

environmental fixed effects, iii) the accounting of all available genetic information however distant genetically or geographically and iv) the provision of expected breeding values for comparison across age groups.

This last point (in conjunction with the first) means that progress can be made by selecting the best individuals of the population no matter what age, no matter how much information is available (individuals with little or no information will only be found at or very close to the mean genetic level of the whole population). It might be advocated that selection should only be from amongst those groups with the highest accuracy. However it is very clearly shown that extra progress is obtained by simulations in pigs (Belansky and Kennedy, 1988) and perhaps more importantly for dairy cattle by the deterministic model of Meuwissen (1989) - in all of these truncation was applied at the same point across all generations irrespective of accuracy. One important consequence is the removal of the requirement to adhere to a rigid strategy that selects individuals at given ages with given accuracies to obtain the required trade-off. Thus the distinction and controversy over MOET and progeny testing disappears. Using BLUP this trade-off is engineered automatically and flexibly. This is shown by results of Meuwissen (1989), given in Table 3. The age distribution uses both young unproven bulls (ages 2-5) and proven progeny-tested bulls (ages 6+). However, since the truncation of BLUP estimates of varying accuracy breaks the conditions for maximum progress to come from the selection of the greatest expected breeding values, BLUP selection may not prove to be optimal.

Table 3 Results of Meuwissen (1989) showing the age distribution (%) of bulls used for breeding in the nucleus and commercial populations, deived from BLUP models

	Age						
Bulls for	2	3	4	5	6	7	8+
Nucleus cows	42	5	3	0	38	9	3
Commercial cows	73	8	3	0	12	3	2

Maximization of progress using BLUP selection

Given the use of BLUP across whole populations, what are the optimal population structures to identify merit as quickly as possible? What is the role of MOET?

Four routes will be explored i) increasing the female reproductive rate, ii) mating design, iii) juvenile predictors, iv) embryo technology.

Embryo Transfer. Meuwissen (1989) has provided a structure and deterministic model of a sufficient complexity to optimise breeding schemes encompassing progeny testing, embryo transfer and nucleus herds of a variable degree of openness. Evaluation in this model mimics BLUP in its use of information and selection across age groups. Initial results (Meuwissen, 1990) suggest that multiple ovulation and embryo transfer in a nucleus herd increases progress by up to 16% in what are termed practical schemes with increases in family size over the range that is currently feasible yielding most of this additional progress.

Mating design. Initial proposals for MOET nucleus schemes used full-sibs primarily because it is the result in practice of obtaining all eggs from a single flush - an ideal that is not always

met. One penalty arising from the production of full-sibs in schemes using family indices is the increased potential for inbreeding through the high correlation of the index amongst them. Nicholas and Smith (1983) attempted to overcome this by restricting bull selection to one per full-sibship thereby reducing the selection proportion n-fold (where n is the number of male full-sibs).

Woolliams (1989b) considered the effect of changing the mating design within the nucleus in order to replace full-sibs with maternal half-sibs. This would be most effectively achieved by mating donor cows to a different bull at each flushing. In ideal circumstances the number of offspring per cow and per bull would remain constant but would result in fewer full-sibs. Table 4 summarises some of the findings. The rested mating design is that used by Nicholas and Smith (1983) with restrictions on the usage of male full-sibs. Results showed that accuracy was marginally reduced by this change. However it was found that greater selection intensities could be applied without incurring additional penalties of inbreeding. Thus for the same rate of inbreeding, expected progress was greater, or alternatively, the same progress could be achieved with less inbreeding. In the absence of selection the mating designs considered by Woolliams (1989b) had identical inbreeding rates, however the imposition of selection created differences in the inbreeding rate. The influence of mating design within a MOET nucleus was not large in the absence of any consideration of inbreeding (including the use of full-sib males) but was relevant to the consideration of progress under constraints on inbreeding.

Table 4 Initial accuracies and rates of progress (phenotypic s.d./annum) together with measures of inbreeding in adult MOET nucleus schemes according to mating design

	L	Accuracy Male	Female	ΔG	Measure of annual inbreeding rate
Nested	3.83	0.56	0.66	0.102	1.73
Factorial	3.83	0.55	0.65	0.117	1.76
	4.18	0.55	0.65	0.107	1.61

The application considered by Woolliams (1989b) was in closed MOET nucleus schemes, however the principles may well be found to apply to the mating design of any nucleus herd employing MOET. The previous section suggests this will still be relevant in population structures that are optimised for BLUP selection.

Juvenile prediction of dairy merit

Woolliams and Smith (1986) considered the impact of a juvenile predictor of dairy merit on genetic progress in conjunction with a variety of breeding schemes. With the computational power available and the amount of performance recording it is only conceivable that such a predictor would be used in conjunction with yield data. When used with a system requiring a progeny test of the bull prior to widespread use the predictor was of value only when combined with MOET. In this case the predictor was used to select which bull from a full-sib family would enter the progeny test. With a co-heritability of 0.27 (an optimistic figure) the predictor gave an additional 10% progress. However the benefits were considerably greater when combined with MOET nucleus schemes (see Table 5) and are

very similar for initial and equilibrium rates of progress (results not shown). The extra progress comes partly from increased accuracy in estimating breeding values of both females and male full sib-ships but approximately half the benefit comes from using the within-family variation among male full-sibs (previously unused) i.e. by offering an informed choice instead of a random one. The value of the predictor was dependent on the accuracy of the scheme in its absence; the value of the extra information being inversely related to the amount of other information, thus juvenile schemes benefitted most. The juvenile predictor has further advantages when inbreeding is a consideration since it shifts the balance of information away from ancestral sources towards the individual's own performance.

Table 5 The benefits of a juvenile predictor expressed as the proportional increase in initial ΔG resulting from its incorporation into indices

h^2		0.1			0.25			0.5		
r_G		0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75
Adult	N*	1.03	1.06	1.11	1.05	1.11	1.19	1.07	1.17	1.31
Juvenile	N	1.06	1.16	1.30	1.09	1.25	1.47	1.13	1.36	1.67

*h^2 = heritability of indicator trait, r_G = additive genetic correlation; + See Table 1

The results of Table 5 are for nested mating designs. The benefits of juvenile predictors in factorial mating systems will not be as dramatic. This is because factorial mating systems improve response by recovering some of the loss of selection intensity that results from the selection among male full-sib families, which is also a part of the improvement given by juvenile predictors. However the additional accuracy from the predictor will still improve progress.

BLUP selection was not considered by Woolliams and Smith (1988). However Table 3 suggests that a very significant proportion of matings would use young unproven bulls. As with factorial mating the benefits of juvenile predictors will depend on the size and usage of full-sib families but additional accuracy will improve progress.

Embryo technologies

Cloning a female dairy cow gives the opportunity to increase the accuracy of evaluation more, and more rapidly, than multiple lactations. Through this route Nicholas and Smith (1983) concluded that cloning would improve rates of progress in MOET nucleus schemes. However their analysis assumed that resources were expandable to accommodate the clone families. Woolliams (1989a) concluded that when resources are fixed the extra accuracy was (in most cases) more than offset by the loss of selection differential through the reduction in the number of distinct genotypes. Exceptions to this were found when the number of sire families as reduced to accommodate the clones but this route would exacerbate inbreeding.

Distinct from enhancing progress, the production of large clone families through nuclear transfer, possibly in combination with embryo stem cells, would produce a genetic lift (a reduction in the time lag between the best of the population and the other sub-populations) through more rapid dissemination of merit. This course is not without risks

which have yet to be quantified. Other benefits include more efficient detection of genotype by environment interactions.

The prospects arising from other embryo technologies were reviewed by Woolliams and Wilmut (1989). They considered in vitro maturation (IVM) and fertilisation (IVF) of oocytes, embryo sexing and freezing, embryo stem cells nuclear transfer (including cloning) and gene transfer. They concluded that embryo sexing, like semen sexing, would be useful in enhancing progress through the more effective use of embryo transfer resources. However although semen sexing has obvious benefits in the conduct of progeny testing, embryo sexing would have practical difficulties that could reduce the robustness of progeny testing. IVM and IVF could improve progress through increasing the female reproductive rate to allow more complex mating designs, however their use will depend greatly upon future improvements in their reliability.

CONCLUSIONS

It is clear from the discussion of Woolliams and Wilmut (1989) and discussions of earlier sections of this paper that the different routes for improving progress are not independent; developments in one route towards increasing progress will sometimes devalue, sometimes enhance developments in another. For example, using embryo sexing in MOET nucleus schemes has obvious advantages in saving resources when only a single member of a full-sibship is used but when factorial mating is introduced, perhaps in combination with a sibship scheme (Ruane and Thompson, these proceedings), or when within-family variation can be utilised using a juvenile predictor, the benefits are reduced.

In conclusion with current technology the maximisation of selection progress in dairy cattle will be brought about through (i) BLUP selection, (ii) the use of MOET to increase the reproductive rate of high-merit cows, and (iii) more attention to mating design to restrict inbreeding. Future developments in physiological and embryological understanding may find a role in increasing progress depending on their reliability and acceptability, and will require careful integration into breeding structures.

REFERENCES

BELONSKY and KENNEDY, B. 1988. J. Anim. Sci. 66 : 1124-1131.
BULMER, M.G. 1980. Mathematical theory of quantitative genetics. Clarendon, Oxford.
DEMPFLE, L. 1974. Genet. Res. 24 : 127-135.
HUDSON, G.F.S. and VAN VLECK, L.D. 1984. J. Dairy Sci. 67 : 161-170.
JUGA, J. and MAKI-TANILA, A. 1987. Act. Agric. Scand. 37 : 511-519.
MEUWISSEN, T.H.E. 1989. Anim. Prod. 49 : 193-202.
MEUWISSEN, T.H.E. 1990. Anim. Prod. (in press).
MEYER, K. and SMITH, C. 1990. Anim. Prod. (in press).
NICHOLAS, F.W. and SMITH, C. 1983. Anim. Prod. 36 : 341-353.
RUANE, J. and THOMPSON, R. 1990. Livest. Prod. Sci. (in press).
SMITH, C. 1969. Anim. Prod. 11 : 433-442.
SMITH, S.P. and HAMMOND, K. 1987. Genet. Sel. Evol. 19 : 181-196.
WOOLLIAMS, J.A. 1989a. Anim. Prod. 48 : 31-35.
WOOLLIAMS, J.A. 1989b. Anim. Prod. 49 : 1-14.
WOOLLIAMS, J.A. and SMITH, C. 1988. Anim. Prod. 46 : 333-345.
WOOLLIAMS, J.A. and WILMUT, I. 1989. Anim. Prod. 48 : 1-30.

Paper 3

Anim. Prod. 1987, 44: 11-19

0003-3561/87/6932001102-00

BIOCHEMICAL AND PHYSIOLOGICAL RESPONSES TO METABOLIC STIMULI IN FRIESIAN CALVES OF DIFFERING GENETIC MERIT FOR MILK PRODUCTION

P. A. SINNETT-SMITH, J. SLEE AND J. A. WOOLLIAMS
†*AFRC Animal Breeding Research Organisation, West Mains Road, Edinburgh EH9 3JQ*

ABSTRACT

Changes in blood metabolite and insulin concentrations in response to metabolic stimuli were evaluated as potential predictors of dairy merit. Calves of high or low genetic merit for milk production were subjected to the stress of : simulated feeds by injection of sodium propionate (0·5 mmol/kg body weight), a short fast and short-term cold exposure. Responses of insulin, free fatty acids, urea, glucose and D-β-hydroxybutyrate were determined by serial blood sampling.

Injections of sodium propionate did not greatly affect blood metabolite concentrations but did sharply increase insulin concentrations; no difference in response between high and low lines was observed.

During fasting serum free fatty acid and D-β-hydroxybutyrate concentrations were greatly increased but no variation due to line was observed. Urea concentrations were also greatly increased during fasting and in this case animals of high dairy merit showed a considerably smaller increase than animals of low dairy merit. Heart rates decreased by one-third during the fast but no variation due to line was apparent; during refeeding high dairy-merit animals had lower heart rates than low-merit animals.

Short-term cold exposure did not significantly alter plasma insulin or metabolite concentrations.

These results are discussed in relation to previous findings and it is concluded that serum urea concentrations during a fast may provide the basis for a useful and robust predictor of dairy merit in young animals.

INTRODUCTION

THE improvement of milk production by genetic selection is limited by the expression of the trait in mature females only. Progeny testing of bulls is used to overcome this limitation, and schemes giving faster genetic progress using multiple ovulation and embryo transfer (Nicholas and Smith, 1983) have been proposed. In all these schemes, further genetic progress could be made by using a measurement in the young bull which is related to breeding value for dairy production.

In lactating cows, genetic differences in milk production appear to be related to variation in metabolic responses to energy demand (Hart, Bines, Balch and Cowie, 1975; Hart, Bines, Morant and Ridley, 1978). Plasma hormone or metabolite concentrations have therefore been investigated as potential predictors of merit but measurements in single, or small numbers of samples from unstimulated animals seem unlikely to provide a useful predictor of breeding value. In consequence, recent studies with young animals have concentrated on responses in animals under some metabolic stress (Tilakaratne, Alliston, Carr, Land and Osmond, 1980; Osmond, Carr, Hinks, Land and Hill, 1981).

In a number of experiments comparing Hereford × Friesian with purebred Friesian calves during fasting or simulated feeding (Land, Carr, Hart, Osmond, Thompson and Tilakaratne, 1983) and purebred dairy calves of high or low genetic merit during fasting (Tilakaratne *et al.*, 1980; Sejrsen, Larsen and Andersen, 1984), differences in hormonal and metabolic responses have been observed. However, the above studies leave important questions unanswered. Could simulated feeds prove an effective discriminator between animals of potentially high and low dairy

† Now called AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station.

merit within a purebred population? Are metabolite responses during fasting generally reproducible in different dairy breed types and sire progeny groups?

In this paper an attempt has been made to answer some of these questions by examining the effects on insulin and metabolite concentrations of (i) energy loading by simulated feeds at various times after food withdrawal (ii) net negative energy balance by fasting and (iii) energy withdrawal by cold exposure, on purebred Friesian calves selected for high or low dairy merit.

MATERIAL AND METHODS

Animals

Purebred British Friesian calves were taken from a herd genetically selected for high or low milk yield. Calves were the progeny of bulls with high or low Improved Contemporary Comparison rating mated to cows that were the products of one to four generations of selection for high or low dairy merit, or to cows of unknown dairy merit. Predicted breeding values, calculated by pedigree analysis are shown in Table 1.

The experiment was carried out in five sessions, between February and August 1984, in order to minimize, as far as practically possible, the spread of ages during experimentation. Additionally, animals in each session were balanced for line and sex. The mean ages and range of ages at the start of each session were 1, 150(132 to 165); 2, 121(110 to 136); 3, 113(108 to 122); 4, 146(130 to 168); 5 126(112 to 152) days.

Experimental procedures

Procedure was the same at each session.

TABLE 1

Predicted breeding values for total milk yield, milk fat and protein (kg per 305-day lactation) of calves, calculated by pedigree analysis (Falconer, 1960)

	High		Low	
	Mean	s.d.	Mean	s.d.
Total yield	557	135	−249	155
Fat	24·6	4·4	−6·5	7·5
Protein	19·1	4·1	−8·0	3·9
No. of calves	17		15	
No. of sires	4		10	

Animals were individually penned for 3 to 4 days prior to the start of the experiment, fed twice daily (08.30 and 16.00 h) with 0·75 kg concentrates (the concentrates consisted of (g/kg): barley 748·5, soya-bean meal 125, salt 5, lime 11, mineral and trace element mix 2·5, calcium dihydrogen phosphate 8 and molasses 100 and had a crude protein concentration of 130 g/kg) with hay provided *ad libitum*. Animals were allowed free access to water throughout. One day before the start, cannulae were placed in both jugular veins, one for blood withdrawal and one for injections. Serum was collected at all sampling times with additional samples being taken for plasma collection using a heparin/fluoride anti-coagulant on specified occasions. Heparinized isotonic saline was used between blood samplings to prevent clotting. Treatment and blood samplings were as follows:

Day 1. Fed as normal; blood sampled at 30-min intervals from 08.30 to 14.30 h.

Day 2. Fed at 08.00; food withdrawn at 08.30 h. Animals were injected with sodium propionate (0·5 mmol/kg body weight, injected over 2 min), 1 h 30 min, 3 and 6 h after withdrawal of food (i.e. 10.00, 11.30, 14.30 h). Blood samples were taken at 08.30, 09.00 h and thereafter at 10-min intervals for 30 min before and 40 min after propionate injections. Additional blood samples for plasma collection were taken at 08.30, 09.00 h, immediately prior to injection and 40 min after injection. No evening feed.

Day 3. Fasted; blood samples taken at 08.30, 12.30 and 16.30 h (for plasma and serum at all times).

Day 4. Sampling started at 07.50 h and was repeated at 10-min intervals until 08.30 h when the animals were re-fed. Thereafter, samples were taken at 09.00, 10.00, 10.30, 11.30, 12.30, 13.30, 14.30 h (for plasma and serum at all times).

Day 5. Animals were fed as normal then a subset of four animals (two high line and two low line) was chosen at random and was subjected to short-term acute cold exposure. Calves were placed in a climate chamber (penned but otherwise unrestrained) maintained at −15°C with air movements of 2 m/sec. Serum and plasma samples were

collected 1 h 15 min, 1 h 25 min, 1 h 35 min and 1 h 45 min after the start of exposure.

Heart rates were measured using electrocardiogram equipment. Recordings were obtained on day 1, 8 h after feeding, on day 2, after completion of propionate injections, on day 3 during fasting and on day 4, 8 h after re-feeding. On each occasion, recordings commenced at least 30 min after blood sampling was completed.

Methods

Assays of urea, glucose, free fatty acids (FFA) and D-β-hydroxybutyrate were carried out using an I.L. Multistat III microcentrifugal analyser (Instrumentation Laboratory (UK) Ltd, Warrington, Cheshire). Commercially available kits were used as instructed by the manufacturers for the analysis of urea (Spinchem, BUN-rate optimized, Smith Kline Instruments Inc., Welwyn Garden City, Hertfordshire) and glucose (Hexokinase method, BCL Ltd, Lewes, East Sussex). FFA concentrations were determined using a WAKO test kit (Alpha Laboratories, Eastleigh, Hampshire) as described by Knox and Jones (1984a). D-β-hydroxybutyrate concentrations were determined as described by Knox and Jones (1984b).

Insulin concentrations in serum were determined by double antibody radioimmunoassay as described by Osmond *et al.* (1981) except that sheep anti-guinea-pig IgG antiserum (Scottish Antibody Production Unit) was used, and the guinea-pig anti-bovine insulin antiserum (Miles Laboratories, Slough) was used at 1:35000 dilution. Samples were assayed in duplicate session by session in a random order and the order of samples within an assay was also completely randomized. The minimum detectable concentration in a single replicate ranged over the assays between 0·16 and 0·37 μg/l). All comparisons of insulin concentrations were made within assays. The coefficient of variation within assays was 0·078.

Statistical analysis

Blood metabolites. Linear functions of the samples for each blood metabolite were calculated for each calf. On day 1, a linear regression on time was fitted by least squares. On day 2, linear functions of the six values taken before and after each propionate injection estimated the average value over all six samples, the linear and quadratic changes in the average value over the three propionate injections, the mean effect of an injection on the metabolite and the linear and quadratic changes in this effect over the three propionate injections. Separately, for day 3 and prior to re-feeding on day 4, mean values were estimated. During the re-feeding period, parameters from linear regression were estimated as for day 1.

Insulin. Concentrations were considered entirely on a logarithmic scale. A baseline level was defined for each animal as the mean of all samples of day 2 excluding those following within 40 min of a propionate injection. The effect of the propionate injection was then assessed, sample by sample, as deviations from this baseline. The mean values on day 3 and prior to re-feeding on day 4 were also assessed as deviations from the baseline. The samples during day 1 and post re-feeding were assessed (i) sample by sample compared to the day 2 baseline, and for post re-feeding, compared to the mean prior to re-feeding and (ii) by linear regression.

The linear functions estimated for each calf, as described above, were analysed on a linear model accounting for sex, line and session with age present as a partial regression. Two-way interactions involving line were investigated but were found not to be significant. Weights accounting for differing numbers of contributing samples between individual calves were used where appropriate.

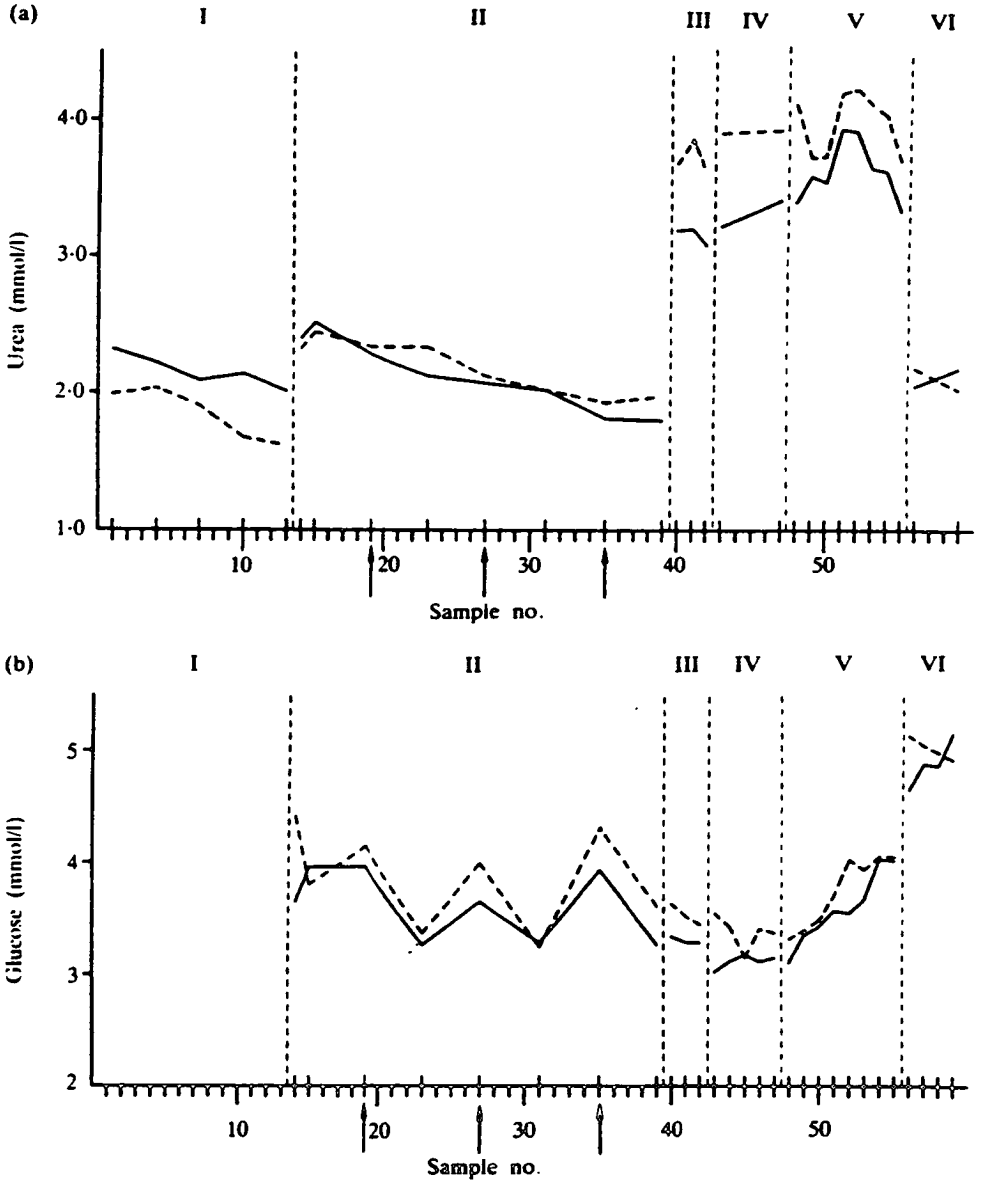
RESULTS

Metabolite concentration changes in response to nutritional challenges

Urea. Baseline urea concentration (mean of day 1) did not show any line or sex differences nor any trends with age although there were some variations between sessions (range 1·51 (s.e. 0·17) to 2·93 (s.e.

0·22) mmol/l). A downward trend with time in urea concentration was observed and there was an indication that the decline was more rapid in low line animals ($P < 0·05$). On day 2, the downward trend in concentration with time was repeated but this was not affected by line. The injections of propionate did not affect urea concentration (Figure 1a).

During fasting, urea concentration increased sharply (Figure 1a). At all times animals of low dairy merit had a significantly higher blood urea concentration than animals of high dairy merit (Figure 1a), the average high-low difference was −0·52 (s.e. 0·25) on day 3 and prior to re-feeding was −0·73 (s.e. 0·23). The divergence between lines varied but this was not related to time, sex, age or session. Upon re-feeding, urea concentration remained high but had decreased to baseline by the start of cold exposure on the following day.



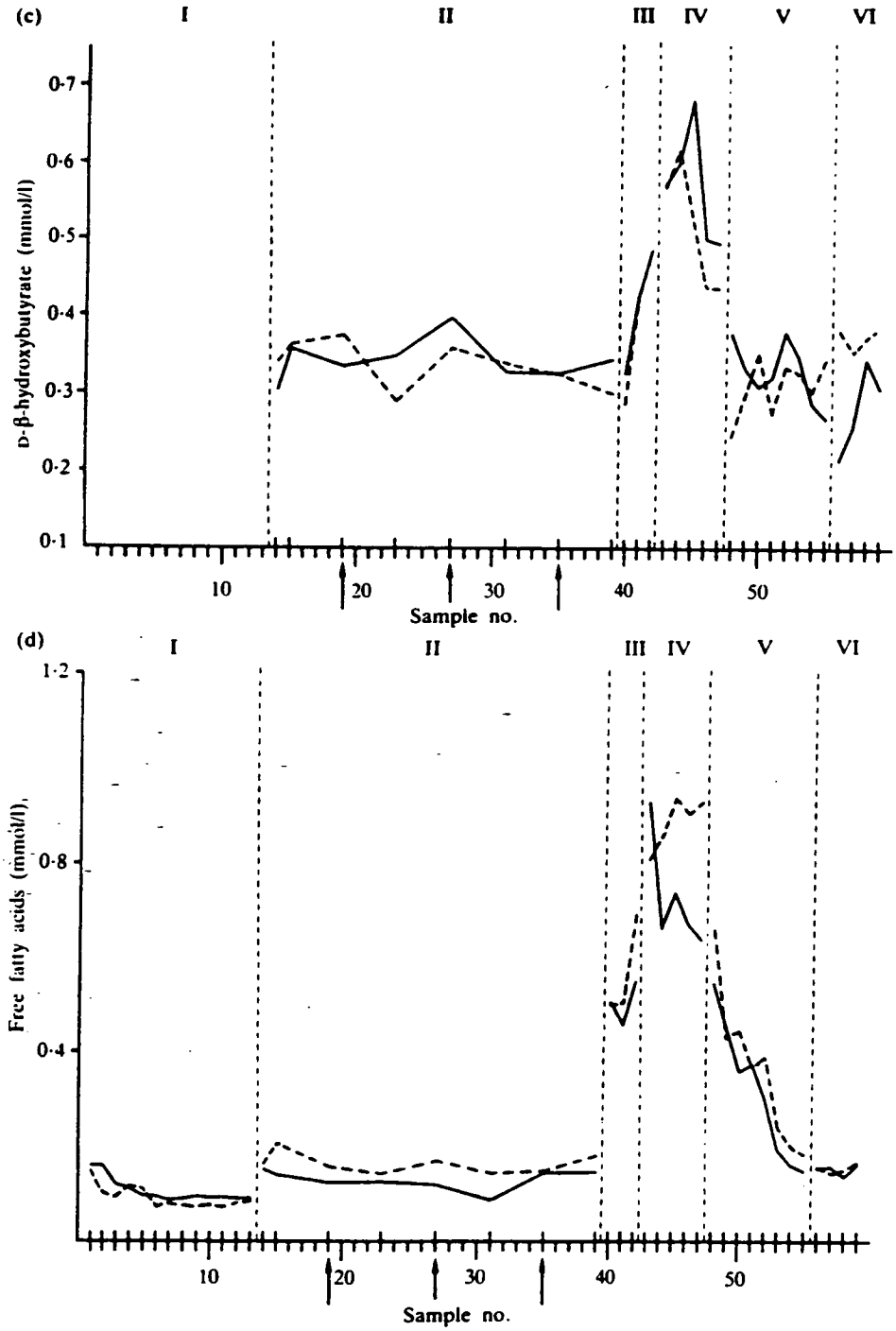


FIG. 1. The blood concentration of (a) urea, (b) glucose (c) D-β-hydroxybutyrate and (d) free fatty acids during day 1 (I) (no treatment) on day 2 (II) in response to propionate injections (↑), during fasting on day 3 (III), immediately prior to refeeding (IV) and after refeeding (V) on day 4, and during cold exposure (VI). Mean concentrations for high (———) or low (- - -) merit animals taken across sessions and sexes. Samples analysed are marked within the axis.

Glucose. Concentration was reduced after propionate injection (Figure 1b) but otherwise largely unchanged by treatment. At no time was there any significant variation attributable to line. A sex difference was observed on day 2, the average value was 3·73 (s.e. 0·104) mmol/l with a male-female difference of 0·694 (s.e. 0·31) mmol/l, and on day 4, session and age differences were apparent with older animals tending to have higher concentrations.

D-β-hydroxybutyrate. There was considerable variation between sessions and concentrations increased during fasting and decreased on re-feeding (Figure 1c). No other significant sources of variation were found.

Free fatty acids. The concentration showed little change during the first 2 days. An increase in concentration due to fasting was observed, which was greater in older animals but no significant line, sex or session differences were detectable. The rate of recovery from fasting did not appear to vary according to line (Figure 1d).

Responses of insulin to metabolic challenges

Insulin concentrations varied considerably over the experimental period (Figure 2). Insulin concentrations have been presented as a baseline value (see MATERIAL AND METHODS) in order to highlight the change; the mean baseline concentration was 0·32 μg/l and did not differ between lines. A downward trend in insulin concentration was observed in high and low lines on day 1, with low line animals having significantly lower insulin concentrations on two occasions ($P < 0·05$; Figure 2) but the mean values for day 1 did not differ.

Insulin concentrations responded rapidly to an injection of sodium propionate and had

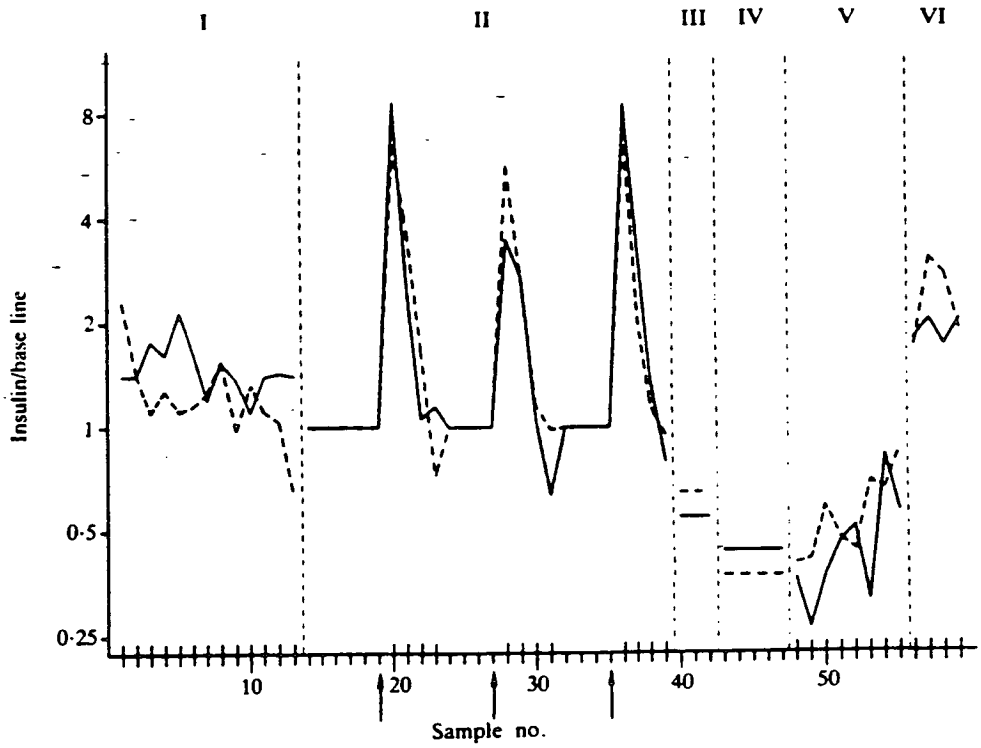


FIG. 2. Serum insulin concentrations during day 1 (I) (no treatment), on day 2 (II) in response to propionate injections (↑), during fasting on day 3 (III), immediately prior to refeeding (IV) and after refeeding (V) on day 4, and during cold exposure (VI). Mean concentrations for high (———) or low (- - -) merit animals, corrected for baseline, taken across sessions and sexes. Samples analysed, and presented individually are marked within the axis, others are means derived as described in MATERIAL AND METHODS.

increased six-fold by 10 min after injection but returned to normal after 40 min. No significant differences in maximal insulin response to propionate were observed. At 20 min after the first injection of propionate, insulin concentrations had started to decrease and the concentration of insulin was closer to baseline in high line animals than in low line animals ($P < 0.05$). The pattern of insulin concentration changes was similar after successive propionate injections but no line differences were observed after the second and third injections (Figure 2). Fasting caused a reduction in insulin concentration which returned to normal upon re-feeding. There was a tendency for insulin concentrations to be lower by the end of the fasting period in the low line but this effect was not significant.

Responses of heart rates to metabolic challenges

The mean heart rate for both high and low animals was lower on the 2nd day than on the 1st day (Table 2) probably due to habituation to handling stress. Fasting greatly decreased heart rates. On re-feeding, heart rates returned to normal in low line animals but remained depressed in high line animals ($P < 0.05$; Table 2).

Responses of metabolites and insulin concentrations to cold exposure

Glucose concentrations were slightly elevated during energy withdrawal by cold exposure but concentrations of urea, D-β-hydroxybutyrate and FFAs were not affected (Figure 1). Insulin concentrations were also elevated after cold exposure (Figure 2). No significant differences between lines were observed.

TABLE 2

Heart rates (beats per min) in different nutritional states

Day	Nutritional state	Mean	s.e.	High-low difference	s.e.d.
1	Fed	80·1	2·7	3·2	4·4
2	After simulated feeding	76·1	2·2	1·1	3·7
3	Fasting	57·4	1·7	−2·6	2·8
4	Re-fed	70·4	3·0	−7·5*	3·2

DISCUSSION

Although sodium propionate injections increased serum insulin concentrations, a line difference was only apparent 20 min after the first propionate injection when concentrations were declining. In a comparison of dairy and beef animals, Land *et al.* (1983) also found differences 20 min after propionate injection, the first sampling after injection. This breed difference (Land *et al.*, 1983) disappeared, as did the increase in insulin following injection, as the time between the last meal and the propionate injection increased from 3 to 48 h. This is in agreement with the present experiment where the only significant line difference was observed following the first injection, 1 h 30 min after the last meal and not following the second and third injections, 3 and 6 h post feeding, although in this study significant increases in serum insulin were still apparent.

Concentrations of urea and FFAs increased greatly during fasting, presumably as a result of amino acid catabolism and fat mobilization respectively. The mobilization of energy reserves is important for high milk production (Broster, Broster and Smith, 1969). The relative responses of urea and FFA concentrations to fasting in Friesian calves of differing merits (Tilakaratne *et al.*, 1980) suggest that animals of high dairy merit derive energy preferentially from fat stores. The present study, and that of Sejrsen *et al.* (1984) in Red Danish calves, confirm that during fasting circulating urea concentrations are lower in calves of high dairy merit. Thus, the urea responses obtained from our work, from Tilakaratne *et al.* (1980) and from Sejrsen *et al.* (1984) are similar despite differences in breed, genetic background and fasting regime. This agreement implies that the line differences in fasting urea concentrations reflect a basic repeatable metabolic difference and not a peculiarity of one breed, line or sire, and confirms the potential of this measurement as a useful predictor. Urea concentration also has the advantage of being easily measured and is relatively resistant to rapid variation due to handling stress.

Our results do not confirm, and those of Sejrsen *et al.* (1984) contradict, the greater

FFA concentrations in animals of high dairy merit observed by Tilakaratne *et al.* (1980). However, these inconsistencies do not rule out the greater use of fat stores as an energy source by high dairy merit animals. Removal of circulating FFAs by utilizing tissues may have differed from experiment to experiment, and may also have been affected by the feeding of straw by Sejrsen *et al.* (1984). This inconsistency of response and the sensitivity of the FFA concentrations to environmental stress (Slee and Halliday 1968) increases experimental variation and underlines the difficulties of using such concentrations as a predictor. It may be that more information could be derived by directly studying adipose tissue lipolysis.

The large drop in heart rate on fasting probably indicates a fall in metabolic rate (Webster 1967). Re-feeding, as expected, returned metabolic rates towards normal. The lower heart rate in high line animals after re-feeding (Table 2) may imply a lower metabolic rate and therefore greater energetic efficiency in these animals in this state.

Cold exposure was intended to provide preliminary data on the effect of short-term energy withdrawal on blood metabolite levels. Webster, Chlumecky and Young (1970) showed that 7-month-old calves, weighing 200 kg, could tolerate temperatures as low as −28°C, but these calves were older and heavier than ours and had an opportunity to become cold-adapted. However, temperatures as high as 7°C were shown to produce a small metabolic response. More specifically the effect of cold on blood metabolites in shorn sheep is such that some increase in serum FFA levels is evident at 2°C whilst exposure to −18°C with wind produces massive elevations associated with hypothermia (Slee and Halliday, 1968). In view of this limited previous information, and also for technical and animal welfare reasons, relatively short 1 h 45 min acute (−15°C) exposures were chosen for the calves in the present experiment. These cold exposures induced shivering but produced no significant change in rectal temperature or in blood metabolite concentrations. It was concluded that longer or more intense cold exposure could be tolerated and would be necessary to produce the desired effects.

In conclusion, this study confirms that blood metabolite concentrations during a metabolic stress in the young animal can relate to its genetic merit for dairy production. In particular, blood urea concentrations during fasting have proved consistent and appear feasible for future use in practice.

ACKNOWLEDGEMENTS

The authors would like to thank Mr A. MacLeod, Mrs C. Woolliams, Miss M. Fordyce, Mr S. B. Wilson and the farm staff at Dryden and Blythbank for assistance. We also thank Dr A. C. Field of the Moredun Research Institute for making analytical equipment available to us.

REFERENCES

Broster, W. H., Broster, V. J. and Smith, T. 1969. Experiments on the nutrition of the dairy heifer. VIII. Effect on milk production of level of feeding at two stages of the lactation. *Journal of Agricultural Science, Cambridge* **72**: 229-245.

Falconer, D. S. 1960. *Introduction to Quantitative Genetics*. Longman, London.

Hart, I. C., Bines, J. A., Balch, C. C. and Cowie, A. T. 1975. Hormone and metabolite differences between lactating beef and dairy cattle. *Life Sciences* **16**: 1285-1292.

Hart, I. C., Bines, J. A., Morant, S. V. and Ridley, J. L. 1978. Endocrine control of energy metabolism in the cow : comparison of the levels of hormones (prolactin, growth hormone, insulin and thyroxine) and metabolites in the plasma of high- and low-yielding cattle at various stages of lactation. *Journal of Endocrinology* **77**: 333-345.

Knox, D. P. and Jones, D. G. 1984a. Automated enzymatic determination of plasma free fatty acids by centrifugal analysis. *Journal of Automatic Chemistry* **6**: 152-154.

Knox, D. P. and Jones, D. G. 1984b. An automated micro-method for the enzymatic determination of D-β-OH butyrate in the ruminant plasma. *Veterinary Research Communications* **8**: 103-115.

Land, R. B., Carr, W. R., Hart, I. C., Osmond, T. J., Thompson, R. and Tilakaratne, N. 1983. Physiological attributes as possible selection criteria for milk production. 3. Plasma hormone concentrations and metabolite and hormonal responses to changes in energy equilibrium. *Animal Production* **37**: 165-178.

Nicholas, F. W. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36**: 341-353.

Osmond, T. J., Carr, W. R., Hinks, C. J. M., Land, R. B. and Hill, W. G. 1981. Physiological attributes as possible selection criteria for milk production. 2. Plasma insulin, tri-iodothyronine and thyroxine in bulls. *Animal Production* **32**: 159-163.

SEJRSEN, K., LARSEN, F. and ANDERSEN, B. B. 1984. Use of plasma hormone and metabolite levels to predict breeding value of young bulls for butterfat production. *Animal Production* **39:** 335-344.

STEE, J. and HALLIDAY, R. 1968. Some effects of cold exposure, nutrition and experimental handling on serum free fatty-acid levels in sheep. *Animal Production* **10:** 67-76.

THAKARATNE, N., ALLISTON, J. C., CARR, W. R., LAND, R. B. and OSMOND, T. J. 1980. Physiological attributes as possible selection criteria for milk production. 1. Study of metabolites in Friesian calves of high or low genetic merit. *Animal Production* **30:** 327-340.

WEBSTER, A. J. F. 1967. Continuous measurement of heart rate as an indicator of the energy expenditure of sheep. *British Journal of Nutrition* **21:** 769-785.

WEBSTER, A. J. F., CHLUMECKY, J. and YOUNG, B. A. 1970. Effects of cold environments on the energy exchanges of young beef cattle. *Canadian Journal of Animal Science* **50:** 89-100.

(*Received* 9 *April* 1986—*Accepted* 9 *July* 1986)

Paper 4

Anim. Prod. 1992, 54: 175-181

0003-3561/92/18460175$02·00

The effect of dietary protein on metabolite concentrations during fasting in calves differing genetically in dairy merit

J. A. Woolliams, R. S. Nisbet and P. Løvendahl†

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

Abstract

Eighty calves of both sexes, 4 months of age, of two lines, differing in genetic merit for fat plus protein production by 70 kg per 305-day lactation, were given one of four diets differing in protein content to give an estimated range of protein intakes from 299 to 530 g/day. After 21 days the calves were fasted for 72 h. Samples of blood were taken before and after feeding in the morning and afternoon prior to fasting, seven further samples were taken throughout the fast and five samples were taken ranging from 0·5 to 7·5 h after food was re-introduced. The samples were analysed for concentrations of urea, free fatty acids (FFA), glucose, triglyceride and β-hydroxybutyrate.

No association between metabolite concentrations and predicted breeding value (PBV) were found at any stage of the trial except for the effect of feeding on FFA concentration, the rate of increase of FFA during the fast and its ultimate concentration. These associations with FFA depended on the sex of the calf and were either absent or less marked in males. The results strengthen the possibility that FFA concentrations during fasting are related to PBV in females, and that sexual dimorphism may explain differing conclusions amongst previous trials.

The magnitude of the estimate of coheritability of urea concentration during fasting with PBV, pooled across experiments, is reduced from –0·193 to –0·066 (s.e. 0·037) by inclusion of these results. However, heterogeneity between experiments is evident and remains to be explained. This suggests that urea concentrations are not consistent predictors of genetic merit and cannot be considered to be reliable for use in practice.

Keywords: *dairy calves, genetic merit, metabolites, protein intake.*

Introduction

The value of an indicator trait that would predict the genetic merit for lactational performance in juvenile male and female cattle was investigated by Woolliams and Smith (1988). They showed that the importance of such a predictor depended on the breeding system within which it was to be incorporated, but that in conjunction with multiple ovulation and embryo transfer the genetic correlation of the indicator trait with yield need only be as great as 0·25 for significant additional genetic progress to be made.

Although the potential value of an indicator was thus clearly established, the problem of identifying such a trait remains. Tilakaratne, Alliston, Carr, Land and Osmond (1980) took the approach that such a trait would most usefully arise from comparing the physiology of high and low merit dairy animals and in particular by investigating the hypothesis of a differential control of energy metabolism. This approach, applied to calves, led them to the finding that the concentration of urea in the blood of calves of genetically high dairy merit had a lower rate of increase during a 48-h fast than in genetically low counterparts. This finding was repeated by Sejrsen, Larsen and Andersen (1984) and Sinnett-Smith, Slee and Woolliams (1987) and the results were summarized by Woolliams and Smith (1988). However, more recently, Mackenzie, Wilson, McCutcheon and Peterson (1988) failed to confirm this result. Nevertheless, the total number of calves investigated in both positive and negative trials was small. Given the importance of the hypothesis more evidence is required.

Present address:
† National Institute of Animal Science, Research Centre Foulum, DK-8830 Tjele, Denmark.

A further problem with the use of metabolite concentrations, and in particular urea, is the influence that environmental factors may have on the result. The concentration of urea in the blood of diverse groups of animals in different environments will depend on many factors but an important consideration is the protein content of their food. Richardson and Kegel (1980) suggested that urea might be used as a measure of protein intake in calves, since a major proportion of the urea found in the blood can originate from protein degradation in the rumen. The rôle of dietary protein prior to fasting in influencing the time course of urea concentrations during a fast, the apparent variation between animals in urea concentrations and its genetic association with yield, all need to be established. This experiment was designed (i) to investigate the effect of dietary protein prior to fasting on urea concentrations during the fast, (ii) to test more rigorously the hypothesis of differential response between groups of calves of differing genetic merit and (iii) to investigate the predictive value of other metabolites.

Material and methods

Animals

Eighty purebred British Friesian calves were sired by one of 14 high (H) merit or one of 11 low (L) merit bulls. Bulls were required to have an Improved Contemporary Comparison (ICC) with a weighting of at least 45 effective daughters. The H and L panels both contributed a total of 40 calves: for H from one to nine calves per bull with the majority between one and three and for L, from one to six calves per bull with the majority between three and five. Genetic information on the remainder of the pedigree varied from unknown at one extreme to known, proven maternal sires for six generations at the other, with the majority between two and four known maternal sires. The predicted breeding value (PBV) of each calf was estimated as the sire's ICC plus, where known, one-half of the ICC of the maternal sire plus one-quarter of the ICC of the maternal grandsire and so on. The predicted difference in the potential yield of fat plus protein between calves of the H and L lines was 70 kg.

Procedure

The calves were born from September 1987 to September 1988 either at Cold Norton Farm, Staffordshire or Blythbank Farm, Tweeddale and were raised on the latter farm from 1 week of age onwards. The calves were given milk-replacer and from 8 weeks onwards were weaned onto hay and concentrate food, this process being complete by 12 weeks of age.

Table 1 *Diets given to calves before and during the experiment*

	Diet			
	1	2	3	4
Composition of concentrate (g/kg)				
Barley	930	825	720	615
Soya-bean meal	0	80	160	240
Fish meal	0	25	50	75
Molasses	50	50	50	50
Vitamins + minerals	20	20	20	20
Energy (MJ/kg DM)	12·63	12·58	12·51	12·48
Crude protein (g/kg DM)	105	150	197	241
Composition of total diet (kg/day)				
Concentrate	2·0	2·0	2·0	2·0
Hay	1·7	1·7	1·7	1·7
Metabolizable energy (MJ/day)	33·5	33·6	33·6	33·6
Crude protein (g/day)	299	375	453	530
Estimated amino acids supplied (g/day)†	250	320	354	387
Proportion of amino acids recommended‡	0·96	1·20	1·33	1·45

† After accounting for rumen degradation (Agricultural Research Council, 1980).
‡ See Agricultural Research Council (1980).

In groups, of either eight or 12, and at an average of 103 days (range 86 to 120 days) calves were individually penned and introduced to one of four diets. All four diets included 1·6 kg hay and 2 kg concentrate per day, but differed in the composition of the concentrate element. Diet 1 contained per kg: 930 g barley, 50 g molasses and 20 g vitamins and minerals; in diets 2, 3, 4 there was a progressive substitution of 105 g barley for a combination of 80 g soya-bean meal and 25 g fish meal. This resulted in four diets of similar energy concentration but in which the protein characteristics differed (Table 1). Diets were given twice daily at 8.30 and 16.00 h and water was freely available. Within 1 week of the introduction of these diets the calves were moved 24 km to Dryden Farm where the subsequent feeding and testing took place.

After a period of 3 weeks there followed a period of 5 days of experimental measurements during which the calves were fasted for a period of 72 h (see Figure 1). At the completion of the fast during which water was freely supplied, the calves were immediately offered half of their daily pre-fasting hay and concentrate diet. However after the first four groups of calves it became clear that this was causing a problem of rumen compaction and so the refeeding schedule was altered so that the initial meal consisted of only a quarter of the daily hay allocation.

The first six groups were balanced as far as possible for line, sex and diet and the following were observed always: (a) for every H calf in a group receiving a given diet there was a L calf (almost always of the same sex) receiving the same diet; and (b) each group contained at least two calves receiving each diet. After group 6 it was decided to give only diet 1 so that (b) no longer applied. Pairs of calves identified by (a) were always penned adjacent to each other.

Sampling

Blood was withdrawn using a needle and syringe on day 1 (the day prior to the start of the fast) at 08.15, 09.00, 15.45 and at 16.30 h (i.e. 15 min before and 30 min after meals); on day 2, at 08.30 (when remaining food was removed) and at 14.30 h (6 h of fasting); on days 3 and 4, at 08.30 and at 14.30 h (24, 30, 48 and 54 h of fasting); and on day 5, at 08.30 (72 h of fasting) and then at 09.00, 10.00, 11.00, 12.00 and 16.00 h (0·5, 1·5, 2·5, 3·5 and 7·5 h) after the reintroduction of food (with the last sample being taken immediately prior to the second meal after fasting finished). The sequence of sampling is shown diagrammatically in Figure 1.

Analytical methods

The blood samples were centrifuged and the serum samples were removed and kept frozen at −20°C. Serum from the first six groups was subsequently assayed using commercial kits (Randox Laboratories Ltd, Crumlin, United Kingdom unless otherwise stated) for concentrations of urea, free fatty acids (FFA; also WAKO, Alpha Laboratories Ltd, Eastleigh), glucose, triglyceride (also Roche Products Ltd, Welwyn Garden City) and β-hydroxybutyrate. For the last two groups, assays for glucose and β-hydroxybutyrate were not carried out.

Statistical methods

The linear model fitted to the results included additive effects of group, sex and diet together with covariate adjustment for age and PBV. The interactions of sex with diet and batch were investigated as were the interactions of group, sex and diet with PBV. However, where interactions were found not to be significant they were deleted from the model. Transformations for the metabolites were investigated using the procedures of Box and Cox (1964). These procedures find the power transformation that has the maximum likelihood under the joint hypotheses of (i) normal errors, (ii) homogenous errors, and (iii) additivity under the model proposed. Logarithmic transformations (which correspond to a power of zero in the family of transformations considered by Box and Cox) were appropriate for urea, FFA and triglyceride concentrations whilst the observed scale i.e. a power

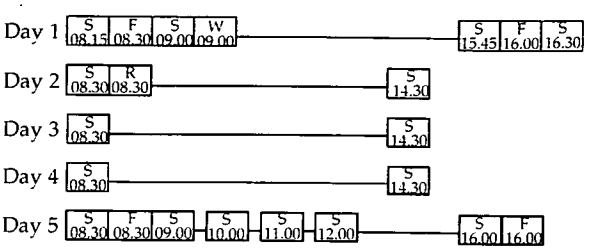


Figure 1 A diagram showing the sequence and timing of sampling of blood, feeding and fasting during the 5-day experimental period. S, sample blood; F, feed; W, weigh; R, remove remaining food.

of 1, was deemed appropriate for glucose and β-hydroxybutyrate.

The data were principally analysed as nine means and contrasts for each metabolite: (a) the mean over day 1; (b) the difference between morning and afternoon samples on day 1; (c) the mean effect of feeding on day 1; (d) the mean over the seven samples taken on days 2, 3 and 4 together with the initial sample on day 5; (e) the linear coefficient with time over this period; (f) the quadratic coefficient; (g) the mean over the samples after refeeding; (h) the linear coefficient with time after refeeding; and (i) the quadratic coefficient. Finally, since primary interest lay in the differential effect of fasting on the metabolites, the mean of the 48-, 54- and 72-h samples and the 72-h sample alone was also analysed by the linear model. The residual effect of sires after accounting for PBV was examined using restricted maximum likelihood (Patterson and Thompson, 1971) but was found not to be of importance.

Results

There were no associations of PBV with metabolite concentrations during the refeeding and further description of results from this period will be omitted.

Urea concentration

There were no differences in response either directly or in combination with other experimental factors attributable to PBV. Over the last 24 h of the fast the urea concentration of H calves was 1·07-fold greater than L calves ($P > 0·05$).

The principal effect on urea concentration was that of diet and Figure 2 shows the pattern of change in urea concentration for each diet. Each increment of dietary protein was associated with an increase in urea concentration during day 1. During this period, the concentrations in calves given diet 1 had increased concentrations in response to feeding.

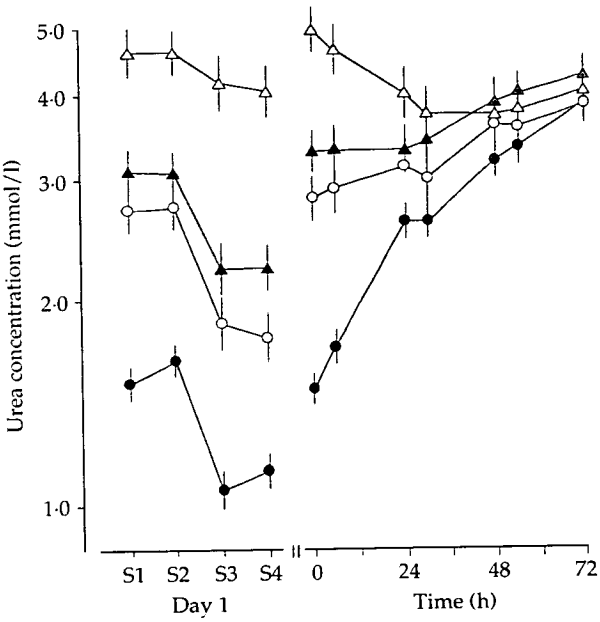


Figure 2 The effect of dietary protein on the concentration of urea during day 1 and throughout fasting. S1, S2, S3 and S4 denote samples taken before and after feeding in the morning and afternoon on day 1. (diet 1 ●; diet 2 ○; diet 3 ▲; diet 4 △).

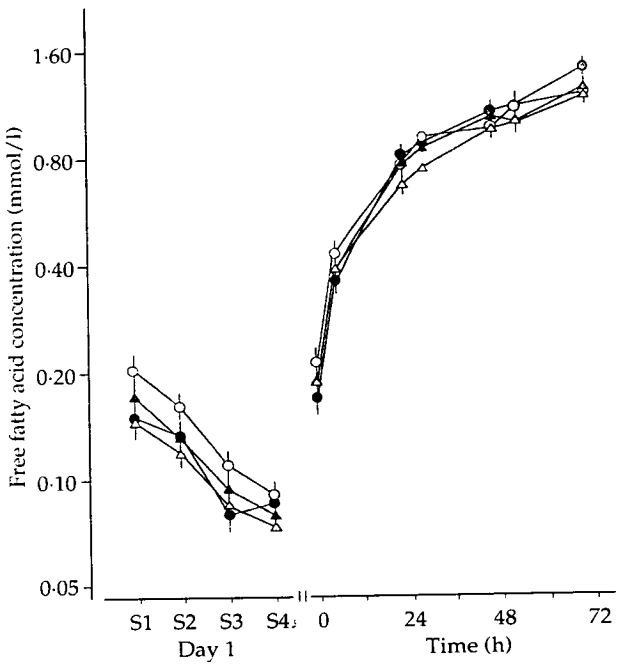


Figure 3 The effect of PBV and sex on the concentration of FFA during day 1 and throughout fasting. S1, S2, S3 and S4 denote samples taken before and after feeding in the morning and afternoon on day 1. (L males △; H males ▲; L females ○; H females ●).

Concentration in calves given the other diets showed little or no response to feeding. The effect of the dietary protein on urea concentrations was evident more than 32 h into the fast. In the course of the fast, concentrations increased in calves given diets 1 to 3 (most rapidly in calves given diet 1) but decreased on diet 4.

The concentrations in male calves were 0·88-fold those of female calves during day 1 ($P < 0·05$) and this proportional difference was slightly reduced to 0·92-fold over the last 24 h of the fast. However the difference in absolute terms (0·32 mmol/l) remained constant.

Free fatty acid concentration

FFA concentrations were associated with PBV in two aspects: firstly in the response to feeding and secondly in the response to fasting. However, in the first and possibly in both of these, the association depended on the sex of the calf. Figure 3 shows the responses according to sex and PBV.

In response to feeding, FFA concentrations in males decreased proportionately to 0·83 of their concentrations before feeding ($P < 0·001$) and calves with high PBV had a marginally greater proportional decrease than those with low PBV: in females, whilst a decrease of a similar order of magnitude was observed in calves of low PBV (0·80-fold, $P < 0·001$) a much smaller decrease was observed in females of high PBV (0·94-fold, $P > 0·05$). This interaction of sex with PBV was significant ($P < 0·01$).

Upon fasting the rate of increase in FFA and the ultimate concentration attained during the fast were positively associated with PBV when the sexes were pooled together ($P < 0·01$ and $P < 0·05$ respectively). Whilst the interaction of PBV with sex was not formally significant ($P > 0·05$) the association in female calves was clearer and more consistent. Since the increase in FFA over the fasting period was clearly non-linear (convex) it is perhaps most useful to examine the mean concentrations over the last 24 h and the final concentration of the fast. On a logarithmic scale the respective regression coefficients on PBV for females alone were 0·0014 (s.e. 0·0009) and 0·0024 (s.e. 0·0010) whilst for males alone the same coefficients were proportionately only 0·51 and 0·34 of these values. As a result, after 72 h fasting, the FFA concentration of H females was 1·18-fold that of L females ($P < 0·05$), and the concentration of H males was only 1·06-fold that of L males ($P > 0·05$).

During day 1, the greater the concentration of protein in the diet, the more nearly constant was FFA maintained from morning to afternoon ($P < 0.001$): thus in calves given diet 1, afternoon concentrations were proportionately only 0·43 of those in the morning whereas in calves given diet 4 this ratio was 0·81 with diet 2 and 3 intermediate. There was no effect of diet on the changes in FFA with fasting.

Triglyceride concentration

The pattern of response is shown in Figure 4. There were no significant associations between triglyceride concentration and PBV. Dietary protein altered the pattern of concentrations on day 1 when an increase in concentration was evident between morning and afternoon in calves given diet 4 ($P < 0.01$), a decrease in calves given diet 1 ($P < 0.05$), with little or no change on diet 2 and 3. At no other stage did dietary protein significantly affect concentrations. During day 1, triglyceride concentrations in males were 0·85-fold those of females ($P < 0.01$) and, whereas concentrations in females tended to increase between morning and afternoon, concentrations in males tended to decrease ($P < 0.05$). The triglyceride concentrations in older calves were more likely to be maintained or increase between morning and afternoon and were initially less likely to decline during fasting. However, by the last 24 h of the fast, there was no effect of age on concentration.

Glucose and β-hydroxybutyrate concentrations

The concentrations of these metabolites were largely unaffected by diet, sex, age or PBV apart from the observation that additional protein in the diet resulted in greater concentrations of glucose during day 1 ($P < 0.05$). Figure 4 shows the pattern of response for both metabolites.

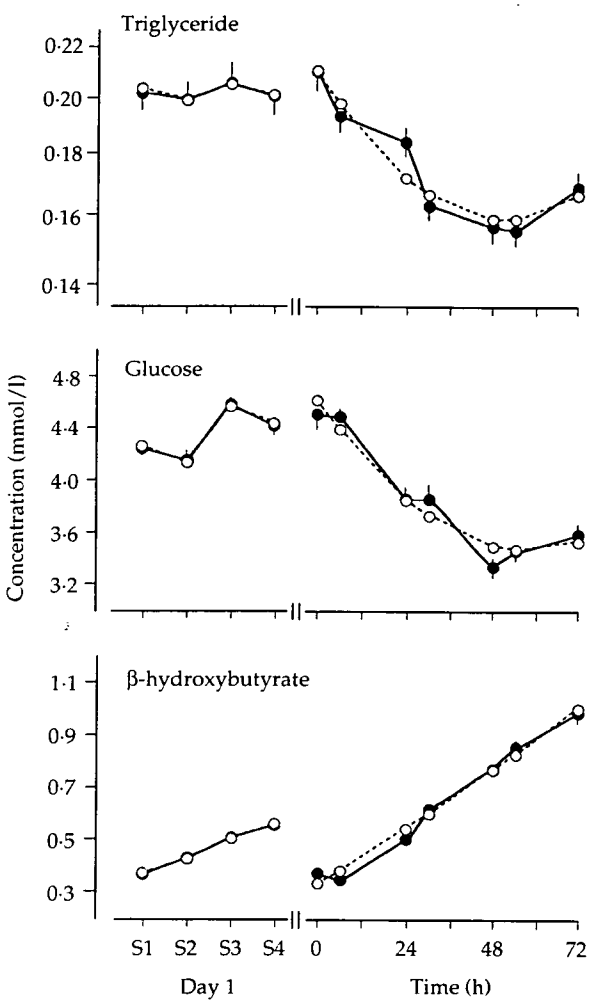


Figure 4 The effect of feeding and fasting on concentrations of (a) triglyceride, (b) glucose and (c) β-hydroxybutyrate, showing fitted means at each sampling point (●——●) and fitted values derived from means and contrasts (O- - -O; see *Statistical methods*). S1, S2, S3 and S4 denote samples taken before and after feeding in the morning and afternoon on day 1.

Discussion

The principle consequence of this study is the re-evaluation of the extent to which a differential response in urea concentrations to fasting can be expected between calves differing in their genetic merit for yield. This is most easily discussed in terms of co-heritability, defined as the regression of the genetic merit for yield on the urea concentration during fasting when both yield and urea are expressed in units of their respective phenotypic standard deviations. In reviewing trial results published to that date (namely Tilakaratne *et al.*, 1980; Sejrsen *et al.*, 1984; Sinnett-Smith *et al.*, 1987), Woolliams and Smith (1988) showed that the coheritability between yield and urea concentration during fasting was consistently negative with an average value of –0·27. The estimate was later revised by Woolliams and Løvendahl (1991) to –0·193 following the results of Mackenzie *et al.* (1988).

The approximate unconditional 95% confidence interval for the coheritability derived from the present trial (distinct from the confidence interval derived directly from the experimental data which is conditional on the genetic sampling of sires; see Woolliams and Smith, 1988) was (–0·053, 0·152) around an expected value of 0·048, a result that is quantitatively distinct from the earlier findings.

Table 2 *A summary of published experimental results concerning the coheritability of urea concentrations in fasted calves with genetic merit for yield*

Reference†	Estimated co-heritability‡	95% confidence interval§	Weighting‖
Tilakaratne *et al.*	−0·26	(−0·53, −0·07)	76
Sejrsen *et al.*	−0·47	(−0·93, −0·22)	32
	−0·17	(−0·47, 0·06)	57
Sinnett-Smith *et al.*	−0·26	(−0·48, −0·09)	105
Mackenzie *et al.*	0·08	(−0·14, 0·33)	72
Present study	0·05	(−0·05, 0·15)	381
Weighted mean −0·066 (s.e. 0·037)			

† See Table 1.
‡ Regression coefficient of breeding value on test result (both measured in phenotypic s.d.) assuming a single sample taken at end of fasting period.
§ Accounts for numbers of calves, numbers of sires and expected genetic divergence in yield (Woolliams and Smith, 1988).
‖ Weighting = [(confidence interval)/4]$^{-2}$.

These results are summarized in Table 2. If weighting factors are derived from the squared length of the approximate unconditional confidence intervals of the other four relevant studies listed above, the present trial has a weight greater than the sum of all others combined. The pooled estimate derived using these weights gives a result of −0·066 (s.e. 0·037). Thus for the first time, the pooled confidence interval includes the value of zero.

There are, however, disturbing discrepancies between the trials such that there cannot be a great amount of confidence even in this pooled estimate. Of the six trials considered (Sejrsen *et al.* (1984) reported on two trials) only two have the pooled estimate within their confidence intervals, and *none* of the three trials with the heaviest weighting has the pooled estimate within their confidence interval. Furthermore, of these three trials, only those of Tilakaratne *et al.* (1980) and Sinnett-Smith *et al.* (1987) have intervals entirely less than the pooled estimate, whilst the present trial has an interval entirely greater than the pooled estimate.

The current trial was designed originally to test for a difference in the power to discriminate merit on diets of differing protein content. Urea concentrations of fed calves have previously been studied with a view to their use in predicting protein content of forage (e.g. Richardson and Kegel, 1980). During day 1, concentrations of urea were up to four-fold higher in calves given the highest protein diet compared with the lowest protein diet. This excess arises principally from degradation of protein in the rumen and the protein concentration of the diet prior to fasting is still a determining factor of urea concentration more than 32 h after the start of fast. However, notwithstanding the importance of the dietary protein, there was no evidence that it affected the ability to predict breeding values for dairy merit at any stage. Thus despite variation in initial pre-fasting concentrations amongst the trials (e.g. Tilakaratne *et al.*, 1980; Mackenzie *et al.*, 1988) dietary protein does not appear to provide an explanation for the heterogeneity of experimental findings.

It appears to the authors that no simple explanation exists for the discrepancies among the various trials. Apart from the dietary differences between trials other explanations could be put forward but could be dismissed equally easily. For example, while Sejrsen *et al.* (1984) and Mackenzie *et al.* (1988) tested only males, there was no suggestion in this trial of a sex difference in the ability to discriminate merit. Another explanation may have been a major gene affecting the response of urea during the fast linked to a major gene for yield segregating in the several populations, from two breeds, that resulted in negative coheritabilities when tested. Yet this possibility and other possibilities based on the sampling of sub-populations are effectively excluded by the fact that this study sampled the same population as Sinnett-Smith *et al.* (1987) only 4 years later but with a comparatively different result. What cannot be ruled out is the effect of chance, but nevertheless if this were the case it would represent an extremely unfortunate pattern of results.

The results concerning the association reported here between FFA and PBV are more encouraging, primarily because they do provide an explanation of heterogeneity between estimates from different trials. In this trial, associations with PBV were found in response to feeding on day 1 and in the ultimate FFA concentration during the fast. However it was also apparent that these associations were either absent or considerably reduced in males. Tilakaratne *et al.* (1980) first reported a positive association of FFA during a fast with PBV but this was not confirmed in the later trials. However, Sejrsen *et al.* (1984) and Mackenzie *et al.* (1988) only tested male calves and so the heterogeneity between trials may be resolved. Neither Tilakaratne *et al.* (1980) nor Sinnett-Smith *et al.* (1987) presented results on sex differences in associations with PBV, and this limits further interpretation. It should be noted that if the association is indeed present in females but absent in males then the value of measuring FFA during a fast as an indicator of genetic merit would be limited (Woolliams and Smith, 1988).

None of the other three metabolites measured appears to be associated with PBV. Mackenzie *et al.* (1988) reported that the control of glucose metabolism was consistently associated with PBV in New Zealand calves from the Massey University experimental herd, an observation based, in part, on similar procedures and measurements to those described here. In view of the fact that other populations have failed to show such an association, it is possible that this observation has arisen through chance genetic associations caused by genetic drift or linkage.

In conclusion, this trial has re-assessed the coheritability of urea during fasting with the result that when pooled across all trials, the estimate is still negative but is no longer distinct from zero. However, unexplained heterogeneity appears to exist amongst the trials contributing to the estimate. Finally, the evidence in female calves for a positive association between FFA concentration at the end of a 48 to 72 h fast and PBV, was strengthened.

Acknowledgements

This experiment was funded by the Milk Marketing Board of England and Wales and the Department of Trade and Industry. The authors are grateful for the help of Mr A. D. Macgregor and his staff at Dryden-Mountmarle Farm, and Mr D. Morrison and Mr W. Lee at Blythbank Farm for their care of the calves; for the essential technical expertise of Mr B. Wilson; and to Dr J. Oldham of the East of Scotland College of Agriculture for his expert advice on the content of the diets.

References

Agricultural Research Council. 1980. *The nutrient requirements of ruminant livestock.* Commonwealth Agricultural Bureaux, Slough.

Box, G. E. P. and Cox, D. R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series B* **26:** 211-252.

Mackenzie, D. D. S., Wilson, G. F., McCutcheon, S. N. and Peterson, S. W. 1988. Plasma metabolite and hormone concentrations as predictors of dairy merit in young Friesian bulls: effect of metabolic challenges and fasting. *Animal Production* **47:** 1-10.

Patterson, H. D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* **58:** 545-554.

Richardson, F. D. and Kegel, L. 1980. The use of biochemical parameters to monitor the nutritional status of ruminants. 1. The relation between concentration of urea nitrogen in the plasma of growing cattle and crude protein content of the diet. *Zimbabwe Journal of Agricultural Research* **18:** 53-64.

Sejrsen, K., Larsen, F. and Andersen, B. B. 1984. Use of plasma hormone and metabolite levels to predict breeding value of young bulls for butterfat production. *Animal Production* **39:** 335-344.

Sinnett-Smith, P. A., Slee, J. and Woolliams, J. A. 1987. Biochemical and physiological responses to metabolic stimuli in Friesian calves of differing genetic merit for milk production. *Animal Production* **44:** 11-19.

Tilakaratne, N., Alliston, J. C., Carr, W. R., Land, R. B. and Osmond, T. J. 1980. Physiological attributes as possible selection criteria for milk production. 1. Study of metabolites in Friesian calves of high or low genetic merit. *Animal Production* **30:** 327-340.

Woolliams, J. A. and Løvendahl, P. 1991. A review of the physiological attributes of male and juvenile cattle differing in genetic merit for milk yield. *Livestock Production Science* **29:** 1-16.

Woolliams, J. A. and Smith, C. 1988. The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46:** 333-345.

(*Received* 29 *April* 1991—*Accepted* 18 *September* 1991)

Paper 5

Response of growth hormone to various doses of growth hormone releasing factor and thyrotropin releasing hormone administered separately and in combination to dairy calves

Peter Løvendahl[1], John A. Woolliams, and Patrick A. Sinnett-Smith

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS, Scotland. *Received 20 Dec. 1990, accepted 31 May 1991.*

Løvendahl, P., Woolliams, J. A. and Sinnett-Smith, P. A. 1991. **Response of growth hormone to various doses of growth hormone releasing factor and thyrotropin releasing hormone administered separately and in combination to dairy calves.** Can. J. Anim. Sci. **71**: 1045–1052. Doses of growth hormone releasing factor (GRF) and thyrotropin releasing hormone (TRH) and combinations of these were administered by intravenous injection to six calves aged 155 ± 3 days and weighing 136 ± 16 kg. Injections were at 09:00, 12:00 and 15:00 h on 4 days, and doses were 0, 15, 30 and 60 pmol GRF kg^{-1} and 0, 275, 550 and 1100 pmol TRH kg^{-1}, with GRF plus TRH at all combinations of these doses. Response of serum growth hormone (GH) was measured as the mean at 5, 10, 15 and 20 min following injection (PEAK) and the area under the curve during 0–60 min (AUC). The correlation between PEAK and AUC was 0.98. The variation in PEAK was related to GH prior to injection and to PEAK 3 h earlier. Separate multiplicative effects for each secretagogue were fitted, with the effects related to the logarithm of dose. Doubling the dose increased PEAK by 1.46-fold following GRF ($P < 0.05$) and 1.25-fold following TRH ($P < 0.05$). There was no evidence that the results for either secretagogue were affected by the presence or absence of the other. This multiplicative model provides a description of the synergy between these secretagogues.

Key words: GH-release, GRF, TRH, calves, dose response

Løvendahl, P., Woolliams, J. A. et Sinnett-Smith, P. A. 1991. **Réponse de l'hormone de croissance (GH) à diverses doses de somatocrinine (GRF) et de thyrolibérine (TRH) administrées séparément ou en association à des veaux laitiers.** Can. J. Anim. Sci. **71**: 1045–1052. Diverses doses, simples et combinées, de GRF hypothalamique et de TRH ont été administrées par voie intraveineuse à six veaux de 155 ± 3 j et pesant 136 ± 16 kg. Les injections étaient faites à 09:00, 12:00 et 15:00 h pendant 4 jours de suite, aux doses de 0, 15, 30 et 60 pmol kg^{-1} pour la GRF et de 0, 275, 550 et 1100 pmol kg^{-1} pour la TRH, en plus de toutes les combinaisons de doses des deux hormones. Pour mesurer la réponse de l'hormone de croissance (GH), on a pris la moyenne des réactions à 5, 10, 15 et 20 minutes après l'injection (PIC), ainsi que la surface comprise sous la courbe de 0 à 60 minutes (AUC). La corrélation entre les valeurs PIC et les valeurs AUC était de 0,98. On a établi les relations entre les variations des valeurs PIC et les valeurs pré-injection, ainsi que les valeurs PIC de l'injection précédente (3 h plus tôt). Les effets multiplicatifs séparés des deux sécrétagogues ont été ajustés en regard du logarithme de la dose. Le doublement de la dose a eu pour effet d'accroître les valeurs PIC de 1,46 fois après injection de GRF ($P < 0,05$) et de 1,25 fois après injection de TRH ($P < 0.05$). Rien n'indique que les résultats observés pour chaque hormone aient été affectés par la présence ou l'absence de l'autre. Ce modèle multiplicatif fournit une description du synergisme des deux sécrétagogues.

Mots clés: Décharge de GH, GRF (somatocrinine), TRH (thyrolibérine), veaux, résponse à la dose

[1]Present address (P.L.): National Institute of Animal Science, Foulum, DK-8830, Tjele, Denmark.

Growth hormone (GH) has important regulatory roles during growth, pregnancy and lactation, but the mechanisms controlling somatotrophs in the pituitary are poorly understood. The release of GH from the anterior lobe of the pituitary is promoted by growth hormone releasing factor (GRF) which in its native form in cattle is a 44 amino acid peptide hormone (Esch et al. 1983; Baile and Buonomo 1987). The sequence of the bovine and the human form shows only one amino acid difference within the first 29 amino acids (Baile and Buonomo 1987). A synthetic peptide containing these 29 amino acids has a GH-releasing potency equal to that of the full-length peptide (Petitclerc et al. 1987).

GH release is also induced in cattle by thyrotropin releasing hormone (TRH) (Johke 1978). Further, a synergistic effect between GRF and TRH on GH release has been reported in cultured bovine pituitary cells (Ingram and Bicknell 1986) and in vivo in lactating cows (Lapierre et al. 1987a) and calves (Hodate et al. 1985), although its physiological importance is unknown. Although this synergy is not affected by photoperiod, it is not evident if GRF and TRH are administered in darkness (Lapierre et al. 1987b).

Dose–response relationships for GRF have been demonstrated in calves (Johke et al. 1984; Della-Fera et al. 1986; Enright et al. 1987), and only limited information is available on the dose response to TRH (Johke 1978; Hedlund et al. 1977). Experiments testing synergy between GRF and TRH have used only single doses of each hormone, thus although the synergistic effect has been established, there is no information on the comparative dose–response relationships of GH release with GRF and TRH in the presence or absence of the synergy. This study investigated such relationships with the long-term objective of obtaining greater understanding of the mechanisms that control the release of GH.

MATERIALS AND METHODS

Treatments and Procedures

ANIMALS. Six British Friesian calves, five females and one male, aged 155 ± 3 days and weighing 136 ± 16 kg were penned individually in a well-ventilated barn with fluorescent tube lights during experimental days. The calves consumed 3–4 kg d^{-1} concentrate (g kg^{-1}: barley 749, soya-bean meal 125, molasses 100, salt and mineral mixture 26) to appetite. Water and hay were available ad libitum throughout.

Synthetic human pancreatic growth hormone releasing factor fragment (1–29) NH_2 (peptide purity > 98%; Bachem, Saffron-Walden, Essex, United Kingdom) was used (henceforth abbreviated to GRF) at doses of 0, 15, 30 and 60 pmol kg^{-1} body weight (denoted G_0, G_1, G_2 and G_3, respectively). Thyrotropin releasing hormone (TRH, Sigma Chemical Company, Poole, Dorset, United Kingdom) was used at doses of 0, 275, 550 and 1100 pmol kg^{-1} body weight (denoted T_0, T_1, T_2 and T_3, respectively). Treatments of GRF plus TRH were at all combinations of these doses. GRF and TRH were dissolved in sterile physiological saline. Saline was also used as vehicle and placebo in a volume of 5 mL per intravenous injection. Twelve of the 16 treatments were allocated to each of the six animals as shown in Table 1.

In the absence of any direct experimental evidence, the decision to give multiple injections to only a few calves was made because previous studies with another pituitary releasing hormone (LH-RH) had shown large variation but with a repeatability of 0.5 (Land 1981). The treatments were given in random order with three injections daily, at 09:00, 12:00 and 15:00 h on four consecutive days.

An indwelling jugular cannula was placed in each animal 1 day prior to the start of the treatments, through which all treatments were given as bolus injections, and blood samples were taken. Samples were taken at −15, −5, 5, 10, 15, 20, 30, 60, 90 and 120 min relative to each injection. The cannulae were kept patent with sterile Na-citrate (4.5% wt vol^{-1}). After centrifugation (2000 × *g*, 4°C, 20 min), serum was harvested and stored at −20°C until assayed.

Table 1. Allocation of calves to treatments

Calf	Treatment[z]
1	(G_0, G_1, G_2, G_3) × (T_0, T_1, T_2)
2	(G_0, G_1, G_2, G_3) × (T_0, T_1, T_3)
3	(G_0, G_1, G_2, G_3) × (T_0, T_2, T_3)
4	(G_0, G_1, G_2) × (T_0, T_1, T_2, T_3)
5	(G_0, G_1, G_3) × (T_0, T_1, T_2, T_3)
6	(G_0, G_2, G_3) × (T_0, T_1, T_2, T_3)

[z](...) × (...) denotes all combinations, e.g., (G_0, G_1) × (T_0, T_1) = (G_0T_0, G_0T_1, G_1T_0, G_1T_1). G_0, G_1, G_2 and G_3 denote GRF at 0, 15, 30 and 60 pmol kg^{-1} body weight. T_0, T_1, T_2 and T_3 denote TRH at 0, 275, 550 and 1100 pmol kg^{-1} body weight.

Assay

GH was assayed using a double antibody radioimmunoassay based on that of Hart et al. 1975 but modified as follows: recombinant DNA-derived bovine GH (American Cyanamid Company, Princeton, NJ) was used for radiolabelling (Iodogen method, Pierce Chemicals, Cambridge, United Kingdom) and standards. Antisera used were guinea-pig anti-bGH and donkey anti-guinea-pig gammaglobulin bound to cellulose (SAC-CEL, Immuno Diagnostic Systems, Washington, Tyne and Wear, United Kingdom).

A 0.05 M borate buffer of pH 7.8 containing 0.2% wt vol^{-1} bovine serum albumin and 0.1% wt vol^{-1} Na azide was used. Serum sample, ^{125}I-bGH and first antibody was incubated 24 h at 4°C. Bound GH was precipitated with SAC-CEL by centrifugation, and ^{125}I-GH bound determined in a gamma counter. At the final dilution of 1 : 40 000 of the first AB in the tubes, binding in the absence of unlabelled hormone was 39%, and the nonspecific binding was 5%. The lowest detectable concentration of GH was between 0.4 and 0.8 ng mL^{-1} standard. Serial dilution of calf serum with buffer paralleled the standard curve. Cross-reactivity with FSH, LH and prolactin was not detectable at concentrations up to 10 μg mL^{-1}. Recovery of known amounts of GH added to serum was 94%. Inter- and intra-assay coefficients of variation were 13 and 6%, respectively (three assays).

Statistical Analyses

The GH concentrations in the single serum samples were combined into the mean of samples taken at −15 and −5 min (PRIOR) and the mean of samples taken at 5, 10, 15, and 20 min after injection (PEAK). Also, the area under the curve of GH concentration plotted against time during the first 60 min after injection (AUC) was calculated. Transformations to natural logarithms were made to obtain approximate normality for all three variables.

Linear mixed models were fitted using routines of GENSTAT (Genstat 5 Committee 1987). Treatments effects were modelled in two analyses: first, a full 4 × 4 factorial (15 df); and second, a reduced model excluding nonlinear effects of the logarithm of dose and their interactions and including factors for GRF given or not (G_1, G_2, G_3) or (G_0); 1 df), for TRH given or not (T_1, T_2, T_3) or (T_0), 1 df) and their interactions (1 df), linear effects of the logarithm of dose of GRF in the absence of TRH (1 df) or in its presence (1 df) and linear effects of the logarithm of dose of TRH in the absence of GRF (1 df) or in its presence (1 df). The between-animal variation was modelled as a blocking factor. Further adjustments were made for days (3 df) and time of injection (2 df). The effect of GH concentration prior to injection was included as a covariate.

From examining the results of the analyses, two further effects required attention. First, despite the logarithmic transformation, considerable heterogeneity in the size of the residual variance was found. One option considered was further transformation, but this would have removed the simple interpretation offered by logarithms. Therefore the heterogeneity was accomodated by application of restricted maximum likelihood (Patterson and Thompson 1971; Davidian and Carroll 1987) and estimating separate residual variances for each of the four subsets of data (G_0,T_0), (G_0) × (T_1, T_2, T_3), (G_1, G_2, G_3) × (T_0) and (G_1, G_2, G_3) × (T_1, T_2, T_3).

Analyses also revealed that carry-over effects were apparent between injections 3 h apart. The validity of the results therefore required that these be adequately modelled. Examination of the residuals showed a strong negative, linear relationship between the residual error and the response obtained 3 h previously (Fig. 1). Therefore the effect was modelled by including the deviation from the mean of the previous response as a covariate. For injections at time 09:00 h this covariate was made equal to zero.

RESULTS

An indication of the time course of GH concentration following intravenous injections is shown in Fig. 2 for the average of the treatments for placebo (G_0) × (T_0), GRF alone (G_1, G_2, G_3) × (T_0), TRH alone (G_0) × (T_1, T_2, T_3) and GRF plus TRH (G_1, G_2, G_3) × (T_1, T_2, T_3). Peak response was observed within 15 min and then declined towards pre-injection concentrations rapidly before 60 min and then more slowly in the next 60 min. The correlation between PEAK over the first 20 min and AUC was high ($r = 0.98$; 70 df; $P < 0.001$), and results will be presented only for PEAK.

The estimated means for each of the 16 treatments after adjustment for the effects described previously are shown in Fig. 3. The difference between the full and reduced models for treatments (see Statistical Analyses) was tested by a likelihood test and was found not to be significant ($P > 0.1$), i.e., the information contained by the data was adequately described by

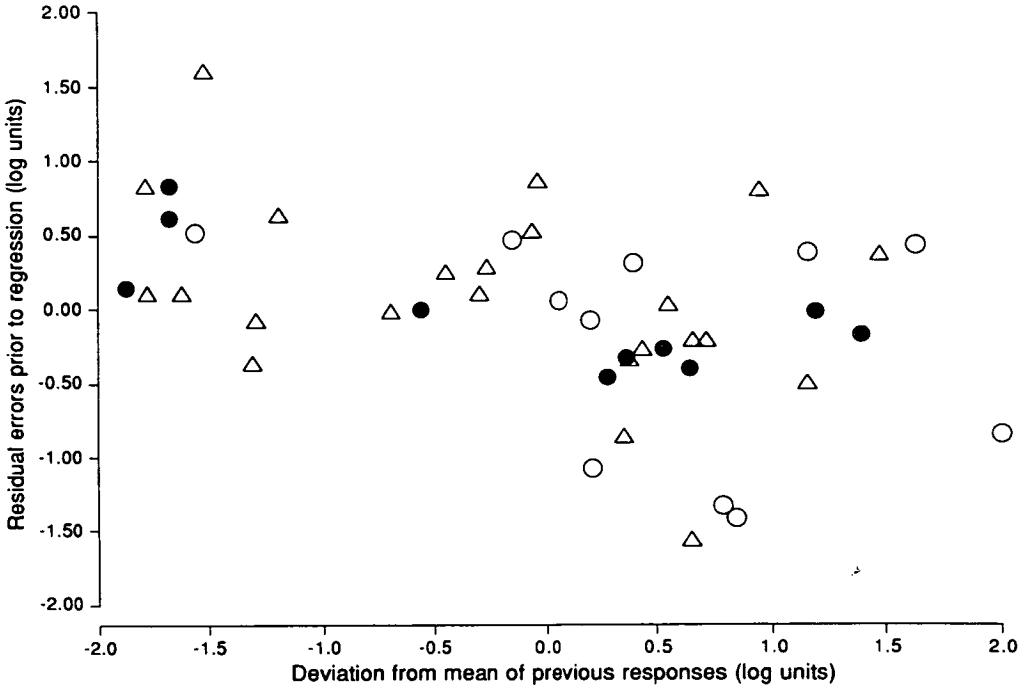


Fig. 1. The effect of magnitude of response 3 h previously on the response following administration of GRF, TRH or GRF plus TRH. (●, TRH alone; ○, GRF alone; △, GRF plus TRH).

the reduced model, involving regressions on the logarithm of doses of each secretagogue.

Using the reduced model, there was no evidence that dose–response relationships for either TRH or GRF differed according to whether each secretagogue was given separately or combined with the other, but the variation associated with administering GRF alone made this dose response the least clear (see Fig. 3). The pooled dose responses were significant for both TRH and GRF over the ranges considered, with mean concentration increasing 1.25-fold ($P < 0.05$) for a doubling of the dose of TRH and 1.46-fold ($P < 0.01$) for a doubling of the dose of GRF.

For GRF alone the lowest dose of 15 pmol kg^{-1} body weight was sufficient to give a detectable response (i..e, significantly greater than saline, with $P < 0.05$) compared with 550 pmol kg^{-1} body weight in TRH alone.

Since, on the logarithmic scale of analysis, the responses were linearly related to the logarithm of dose (interaction terms over and above the reduced model were not significant), it is possible to obtain further information on the relationship between the secretagogues by comparing the average of the treatment means for GRF alone (G_1, G_2, G_3) × (T_0), TRH alone (G_0) × (T_1, T_2, T_3,) and GRF plus TRH (G_1, G_2, G_3) × (T_1, T_2, T_3) with saline (G_0, T_0). Following injection of GRF alone PEAK was 4.44-fold greater ($P < 0.001$) than that following saline. Following TRH alone, PEAK was only 1.55-fold greater ($P < 0.001$). PEAK following GRF plus TRH was 9.22-fold greater ($P < 0.001$) than following saline. On the scale of analysis there was no evidence of synergism ($P > 0.1$), since the multiplicative effect was consistent with the product of separate contributions from GRF and TRH (i.e., 9.22/(4.44 × 1.55) = 1.34, and the estimate of 1.34 was not significantly different from 1).

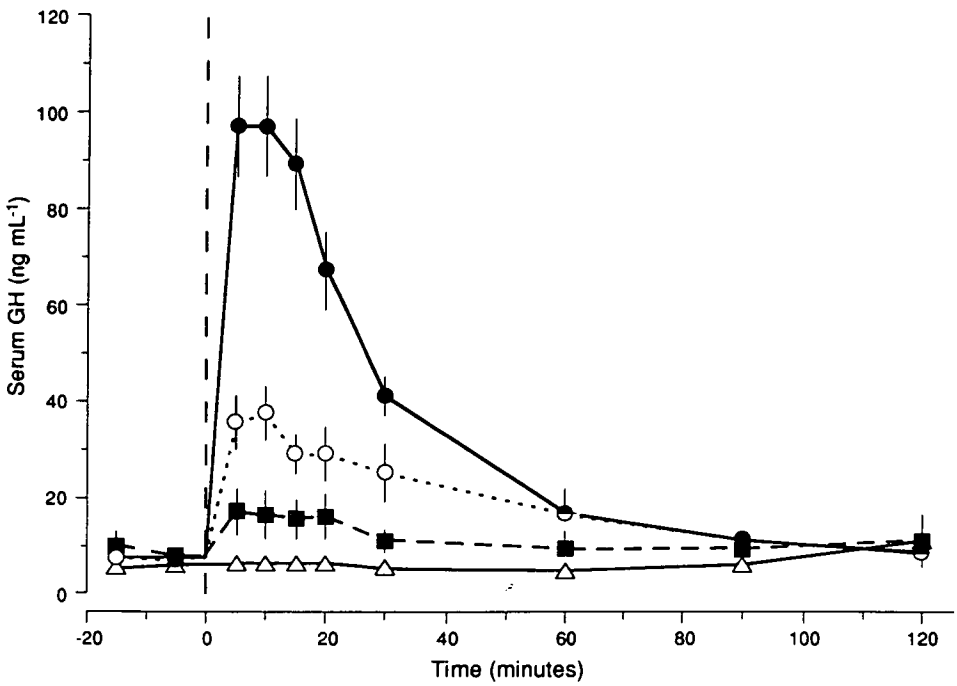


Fig. 2. Serum GH response in calves following administration of saline (▵ , n = 6 trials), GRF pooled over doses of 15, 30 and 60 pmol kg^{-1} body weight (○ , n = 15 trials), TRH pooled over doses of 275, 550 and 1100 pmol kg^{-1} body weight (■ , n = 15 trials) and combined administration of GRF and TRH pooled over the same doses (● , n = 36 trials). Values are mean ± SE.

PRIOR was not affected by any of the experimental factors. However, PEAK was associated with PRIOR, and the partial regression coefficient of PEAK on PRIOR was 0.78 ± 0.18 ($P < 0.01$). Since this was on logarithmic scales, it suggests a doubling of PRIOR would be associated with a 1.72-fold increase in PEAK. This relationship did not differ according to whether GRF, or TRH or both were administered.

The relationship between response and the response to a treatment 3 h previously is shown in Fig. 1. The relationship was modelled by linear regression, with a coefficient of −0.31 ± 0.07 ($P < 0.01$).

Variation unexplained by the between-animal variation or the fixed effects was large, and heterogeneity was marked even after transformation onto the logarithmic scale. After adjusting for the models described, the residual variances ranked "saline," treatments involving "TRH alone," treatments involving "GRF plus TRH" and treatments involving "GRF alone," with associated coefficients of variation on the observed scale of 33, 38, 69 and 83%, respectively. Thus administration of GRF resulted in greater coefficients of variation. This variation was a major contributor to the difficulty encountered in establishing a clear dose response for GRF when administered on its own.

DISCUSSION

To study GH responsiveness in dairy calves it is necessary to have clear information on dose-response curves of potentially important secretagogues such as GRF and TRH, including minimum doses, the position of the curve most sensitive to dose. Further, in the absence of prior evidence on the physiological importance of synergy between GRF and TRH, information is needed on the two-dimensional joint response curve for TRH and GRF.

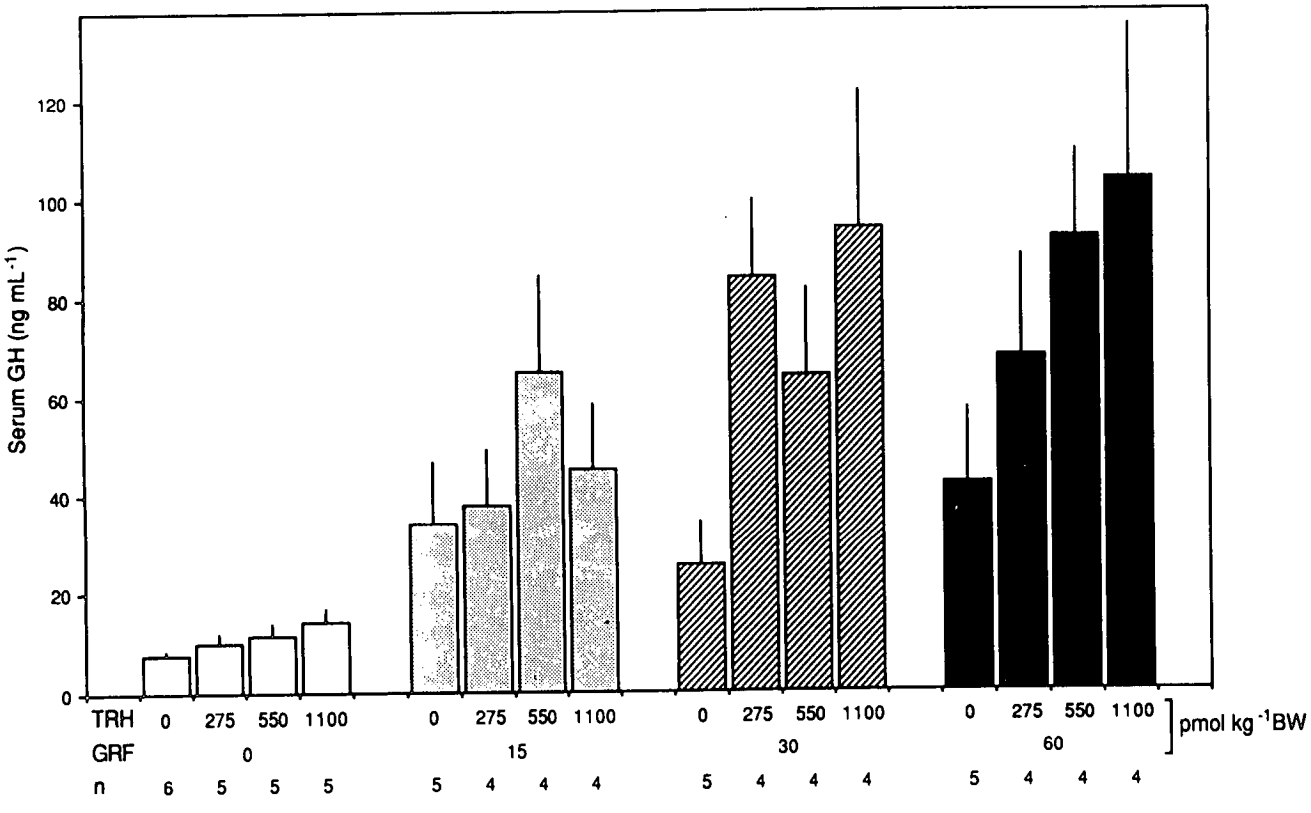


Fig. 3. Mean concentration of GH for 20 min following the intravenous injection of GRF (0, 15, 30 and 60 pmol kg $^{-1}$ BW) and TRH (0, 275, 550 and 1100 pmol kg $^{-1}$ BW), and all combinations of GRF and TRH (unshaded, (T_0, T_1, T_2, T_3); shaded, (T_0, T_1, T_2, T_3) × (G_1); diagonally shaded, (T_0, T_1, T_2, T_3) × (G_2); and solid, (T_0, T_1, T_2, T_3) × (G_3). Values are mean ± SE.

This trial has provided information on these aspects beyond that previously published and has clearly identified factors important to the precision of subsequent trials.

The relationship between response and GH concentration prior to injection for both GRF and TRH and the increased coefficient of within-animal variation when GRF was injected were both important in determining precision and may both be related to the synchrony of endogenous GH-pulsing and exogenous stimuli. Tannenbaum and Ling (1984) showed in rats that GH response to GRF was very small during an endogenous GH trough compared to when GRF was given during an endogenous GH peak. For both secretagogues, size of response may depend in part on the size of the releasable GH pool in the pituitary, and this would be affected by the time since the most recent peak. Further, the response to GRF may also be more profoundly affected by the endogenous somatostatin. Thus regression on GH concentrations prior to injection gave an improvement in precision, but variation in response to GRF remained large.

A further factor influencing response, and hence precision, was found to be the carry-over effect from injections 3 h previously. This is despite two studies in calves (Plouzek et al. 1983; Hodate et al. 1985) which have used repeated injections at 3-h intervals and have explicitly reported no such effects and other studies using slightly longer intervals (e.g., Enright et al. 1987) in which carry-over effects were not described. Nevertheless, the finding in this trial was clear, and a negative linear relationship was found between the magnitudes of responses 3 h apart. Although it was possible to model the carry-over effect, longer time intervals between injections should be used to avoid such a problem altogether.

The present trial showed that the threshold dose for an observable GH release in calves was less than 15 pmol GRF kg^{-1} body weight. Upper bounds to this dose have been variously estimated with different fragments as 6.5 pmol kg^{-1} (GRF(1-40) Della-Fera et al. 1986), 36 pmol kg^{-1} (GFR(1-44), Moseley et al. 1984), 90 pmol kg^{-1} (GRF(1-40), Enright et al. 1987) and 67 pmol kg^{-1} (GRF(1−29) and GRF(1-40), Petitclerc et al. 1987). The generality of these results relies on the equipotency of fragments shown by Petitclerc et al. (1987). The result of this trial places a firm emphasis on the lowest estimate of 6.5 pmol kg^{-1}. The variation in response and failure to cope with heterogeneity of data are probable causes of the high estimates in some of the studies mentioned in which lower doses were deemed not to have elicited responses.

Dose responsiveness for GRF was observed between 15 and 60 pmol kg^{-1}, although the variation made this difficult to establish, and taken together with the dose–response studies mentioned previously, suggests the range of doses used in this trial contains the most sensitive region of dose response.

For TRH, the threshold derived from this trial lies between 275 and 550 pmol kg^{-1} bodyweight. Previous estimates were less than 300 pmol kg^{-1} in cows (Johke 1978) and less than 325 pmol kg^{-1} in calves (Tucker and Wettemann 1976). These results are consistent and suggest that the lowest dose used here is very near to the threshold. Hedlund et al. (1977) gave TRH to calves at doses greater than those used here and showed that compared with GRF the maximal response is much smaller and that little responsiveness is observed above 1375 pmol kg^{-1}. Thus a tentative conclusion is that the highest dose used in this study, 1100 pmol kg^{-1}, is close to the ED_{50}.

In the regions of the dose-response curve examined, effects of TRH and GRF appeared multiplicative and so exhibited little interdependence on a logarithmic scale. There was no evidence of different regressions of the logarithm of response on the logarithm of dose for one secretagogue in the presence or absence of the other, and the comparison of treatments with GRF alone, TRH alone and GRF plus TRH also suggested this. Thus in the absence of precise mechanisms and dynamic equations to describe the interaction, the synergy reported on the observed scale (Hodate et al. 1985) may be reduced by considering the effects on a logarithmic scale.

In conclusion, the results, in characterising more fully the two-dimensional dose–response surface for TRH and GRF and in identifying factors influencing the large variation in response, will allow better design and more informed interpretation of experiments to understand the pituitary responsiveness to GH secretagogues in calves.

ACKNOWLEDGMENTS

The authors would like to thank Mr. B. Wilson for technical assistance and Mrs. J. Goode (Institute of Animal Physiology and Genetics, Babraham, United Kingdom) for help with development of the assay. This work was supported by the Underwood Fund (AFRC) and the Danish Research Academy.

Baile, C. A. and Buonomo, F. C. 1987. Growth hormone-releasing factor effects on pituitary function, growth, and lactation. J. Dairy Sci. **70**: 467–473.

Davidian, M. and Carroll, 1987. Variance function estimation. J. Am. Stat. Assoc. **82**: 1079–1091.

Della-Fera, M. A., Buonomo, F. C. and Baile, C. A. 1986. Growth hormone releasing factors and secretion of growth hormone in sheep, calves and pigs. Domest. Anim. Endocrinol. **3**: 165–176.

Enright, W. J., Zinn, S. A., Chapin, L. T. and Tucker, H. A. 1987. Growth hormore response of bull calves to growth hormone-releasing factor. Proc. Soc. Exp. Biol. Med. **184**: 483–488.

Esch, F. S., Bohlen, P., Ling, N. C., Brazeau, P. E. and Guillemin, R. 1983. Isolation and characterization of the bovine hypothalamic growth hormone releasing factor. Biochem. Biophys. Res. Comm. **117**: 772–779.

Genstat 5 Committee. 1987. GENSTAT 5 reference manual. Clarendon Press, Oxford, United Kingdom.

Hart, I. C., Flux, D. S., Andrews, P. and McNeilly, A.S. 1975. Radioimmunoassay for ovine and caprine growth hormone: its application to the measurement of basal circulating levels of growth hormone in the goat. Horm. Metab. Res. **7**: 35–40.

Hedlund, L., Doelger, S. G., Tollerton, A. J., Lischko, M. M. and Johnson, H. D. 1977. Plasma growth hormone concentrations after cerebroventricular and jugular injection of thyrotropin-releasing hormone. Proc. Soc. Exp. Biol. Med. **156**: 422–425.

Hodate, K., Johke, T. and Ohashi, S. 1985. Growth hormone, thyrotropin and prolactin response to simultaneous administration of human growth hormone-releasing factor and thyrotropin releasing hormone in the bovine. Endocrinol. Jpn. **32**: 375–383.

Ingram, C. D. and Bicknell, R. J. 1986. Synergistic interaction in bovine pituitary cultures between growth hormone-releasing factor and other hypophysiotrophic factors. J. Endocrinol. **109**: 67–74.

Johke, T. 1978. Effects of TRH on circulating growth hormone, prolactin and triiodothyronine levels in the bovine. Endocrinol. Jpn. **25**: 19–26.

Johke, T., Hodate, K., Ohashi, S., Shiraki, M. and Sawano, S. 1984. Growth hormone response to human pancreatic growth hormone-releasing factor in cattle. Endocrinol. Jpn. **31**: 55–61.

Land, R. B. 1981. Physiological criteria and genetic selection. Livest. Prod. Sci. **8**: 203–213.

Lapierre, H., Petitclerc, D., Dubreuil, P., Pelletier, G., Gaudreau, P., Morisset, J., Couture, Y. and Brazeau, P. 1987a. Synergism between a human growth hormone-releasing factor and thyrotropin-releasing factor on growth hormone release in relation with the stage of lactation of Holstein dairy cows. Reprod. Nutr. Dev. **27**: 605–607.

Lapierre, H., Petitclerc, D., Pelletier, G., Dubreuil, P., Morisset, J., Gaudreau, P., Couture, Y. and Brazeau, P. 1987b. Synergism and diurnal variations of human growth hormone-releasing factor (1–29)NH_2 and thyrotropin-releasing factor on growth hormone release in dairy calves. Domest. Anim. Endocrinol. **4**: 207–214.

Moseley, W. M., Krabill, L. F., Friedman, A. R. and Olsen, R. F. 1984. Growth hormone response of steers injected with synthetic human pancreatic growth hormone-releasing factors. J. Anim. Sci. **58**: 430–435.

Patterson, H. D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika **58**: 545–554.

Petitclerc, D., Pelletier, G., Lapierre, H., Gaudreau, P., Couture, Y., Dubreuil, P., Morisset, J. and Brazeau, P. 1987. Dose response of two synthetic human growth hormone-releasing factors on growth hormone release in heifers and pigs. J. Anim. Sci **65**: 996–1005.

Plouzek, C. A., Anderson, L. L., Hard, D. L., Molina, J. R. and Trenkle, A. 1983. Effect of intravenous injection of a growth hormone releasing factor on plasma growth hormone concentrations in cattle. J. Anim. Sci. **57**(Suppl.): 203.

Tannenbaum, G. S. and Ling, N. 1984. The interrelationships of growth hormone (GH)-releasing factor and somatostatin in generation of the ultradian rhythm of GH secretion. Endocrinology **115**: 1925–1957.

Tucker, H. A. and Wettemann, R. P. 1976. Effects of ambient temperature and relative humidity on serum prolactin and growth hormone in heifers. Proc. Soc. Exp. Biol. Med. **151**: 623–626.

Paper 6

The effect of genetic selection for milk yield on the response to growth hormone secretagogues in immature cattle

P. Løvendahl, K. D. Angus and **J. A. Woolliams**

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

(P. Løvendahl was on leave from National Institute of Animal Science, Foulum, DK-8830, Tjele, Denmark)
(Requests for offprints should be addressed to J. A. Woolliams)

RECEIVED 18 June 1990

ABSTRACT

Eighty 4-month-old calves of both sexes and of two selected lines differing by 70 kg in their predicted total yield of milk fat and protein were injected intravenously with three of four GH secretagogues: these were, per kg liveweight; (i) 0·2 μg human GH-releasing factor(1–29) (GRF), (ii) 0·2 μg TRH, (iii) a combination of (i) and (ii), and (iv) 0·1 g arginine hydrochloride. The response of GH was measured for 2 h following administration. Geometric mean concentration of the 5-, 10-, 15- and 20-min samples following GRF, TRH and their combination were 29·3, 19·5 and 156 μg/l compared with baseline means of 6·5, 10·0 and 12·6 μg/l respectively, and for arginine (in which the mean response included the 30-min instead of the 5-min sample) 14·6 μg/l compared with a baseline of 8·31 μg/l. The line selected for greater yield responded more to each secretagogue by 1·53-fold following GRF ($P<0·01$), 1·34-fold following TRH ($P<0·05$), 1·11-fold following the combination ($P>0·01$) and 1·26-fold following arginine ($P<0·1$). Females responded 2·3-fold more than males following GRF administration ($P<0·001$), only 1·2-fold more following TRH ($P>0·1$), but less (0·63-fold) than males when GRF was combined with TRH ($P<0·05$). For all secretagogues the concentration of GH before administration was important in determining the size of response ($P<0·001$). It was concluded that the increased release of GH following the administration of GRF and TRH was a direct result of selection for dairy merit and that increased yields during lactation may, in part, be mediated directly through pituitary responsiveness.
Journal of Endocrinology (1991) **128,** 419–424

INTRODUCTION

Obstacles to progress in dairy cattle breeding are first that milk yield is only expressed in mature females and secondly that the reproductive capacity of females is low compared with that of males. Indicator traits for milk yield, measurable in juvenile cattle of both sexes, could improve the annual genetic progress, especially when combined with breeding schemes using multiple ovulation and embryo transfer (Woolliams & Smith, 1988).

An important difference between high and low yielding cows is the priorities given to the mammary gland and the body tissues in the distribution of nutrients. Associations between milk yield and plasma concentrations of growth hormone (GH), both temporally throughout lactation, and between cows, suggest a lactotrophic effect of this hormone in cattle (Hart, Bines, Morant & Ridley, 1978). This suggestion is supported by the fact that administration of exogenous GH of pituitary or recombinant origin consistently increases milk yield some 10–20% (e.g. Chilliard, 1989). Exogenous GH mediates changes in metabolism resembling those brought about by selection for increased milk yield (Peel & Bauman, 1987). These findings suggest that variations in endogenous GH secretion may mediate genetic differences in milk yield. However, the association of endogenous GH with milk yield in lactating cows is confounded by concurrent associations with net energy balance.

Growth hormone secretion is measurable in cattle of both sexes and before sexual maturity. The measurement of GH secretion requires frequent blood sampling because of its pulsatile secretion pattern (Anfinson, Davis, Christian & Everson, 1975). However, some information on pituitary responsiveness

  Printed in Great Britain
0022–0795/91/0128–0419 $02.00/0

can be obtained by the administration of exogenous secretagogues, a method often used in diagnosing GH deficiency (e.g. Takano, Hizuka, Shizume *et al.* 1984).

This experiment was designed to investigate the endogenous GH release of dairy calves differing in their genetic merit for milk yield. Growth hormone secretion was induced by acute intravenous administration of GH-releasing factor, thyrotrophin-releasing hormone and arginine hydrochloride, into immature calves of both sexes.

MATERIALS AND METHODS

Animals and housing

A total of 80 British Friesian calves of both sexes (41 females and 39 males) and of a mean age of 124 days (range, 107 to 141 days) were used in the experiment, 40 from each of two lines selected for either high or low milk yield. Milk yield was measured by the total yield (kg) of fat and protein during a single lactation and individual calves were assigned a predicted breeding value (PBV; a measure of their genetic merit for milk yield) calculated according to the previously established breeding value of male ancestors used during the selection. The mean PBVs were −25·9 (range, −45·8 to 3·1) and 45·8 (range, 27·0 to 62·9) for the low- and high-yield lines respectively. Twenty-five paternal half-sib groups were represented, varying in size from one to nine calves. The mean weight of calves was 116 and 123 kg for females and males respectively with no difference evident between the lines.

During the experiment calves were housed in individual pens in a barn with artificial light providing 16 h light and 8 h darkness. The calves were trained to wear halters in anticipation of their being tethered during sessions of blood sampling.

Experimental design

The experiment was conducted using groups of eight or twelve calves over the period of 1 year. On day 1 of the experiment the calves were housed and introduced to one of four rations, each consisting of 2 kg concentrates and 1·7 kg hay and supplying 335 MJ/day. The concentrate portion of diet 1 consisted (per kg) of 930 g barley, 50 g molasses and 20 g vitamins and minerals. Diets 2, 3 and 4 were produced by progressively substituting 105 g barley with 80 g soya-bean meal and 25 g fishmeal. This resulted in an estimated supply of amino acids beyond the rumen of 250, 320, 354 and 387 g/day, equivalent to 0·96, 1·2, 1·33 and 1·45 times the recommendations of the Agricultural Research Council (1980). The daily ration was fed in two equal feeds at 08.30 h and 16.00 h. On day 22 the calves were weighed and from day 23 they were fasted for 72 h before returning to their respective diets. Water was freely available at all times.

Groups of calves were balanced as far as possible for selection line, sex and diet, and the following protocols were always observed: (a) for every calf of high PBV in a group receiving a given diet there was a calf of low PBV (almost always of the same sex) receiving the same diet, and (b) each group contained at least two calves receiving each diet. After group 6 it was decided to feed only diet 1 so that (b) no longer applied. Pairs of calves identified by (a) were always penned adjacent to each other. Otherwise allocation to pens was at random.

On day 30 a cannula was inserted into the jugular vein and kept patent using sodium citrate (4·5%, w/v). On days 31 and 32 each group of calves was given three of four GH secretagogues in bolus doses. The four secretagogues and doses (per kg liveweight on day 22) were: (i) 0·2 μg human pancreatic GH-releasing factor fragment (GRF(1–29)-NH_2; Bachem, Saffron Waldon, Essex, U.K.), (ii) 0·2 μg thyrotrophin-releasing hormone (TRH; Sigma Chemical Co. Ltd, Poole, Dorset, U.K.), (iii) a combination of (i) and (ii) and (iv) 0·1 g arginine hydrochloride (Sigma Chemical Co. Ltd). Sterile saline was used to dissolve GRF and TRH in a total injection volume of 5 ml. The arginine hydrochloride was dissolved in distilled water (30%, w/v) and neutralized with NaOH to pH 7·4; as a result the injection volume varied according to weight.

The secretagogues were administered and blood was sampled through the indwelling cannulae. Blood samples were taken at −15, −5, 5, 10, 15, 20, 30, 60, 90 and 120 min relative to administration. Serum was harvested by centrifugation, and kept frozen (−20 °C) until assayed for GH. The injections were given at 09.00 h and 12.00 h on day 31 and at 09.00 h on day 32, with all calves in each group receiving the same secretagogues at the same time. Table 1 shows the secretagogues administered and their order according to group. Arginine was administered on day 32 to allow overnight fasting by omitting the p.m. and a.m. feeds on days 31 and 32 respectively, as previous evidence (McAtee & Trenkle, 1971) indicated that this procedure would increase the GH secretory response. Ambient temperature and light intensity were measured immediately before the administration of each secretagogue. Light intensity was measured from 13 standard positions within the barn.

Serum was assayed for GH using a modified form of the radioimmunoassay of Hart, Flux, Andrews & McNeilly (1975). Briefly, the assay used a guinea-pig bovine GH antibody at a final dilution 1:40 000, in a borate buffer (0·05 mol/l; pH 7·8). Samples were incubated with the first antibody and tracer (15 000 c.p.m.)

TABLE 1. The order of secretagogue administration according to group

	No. of calves	Day 31		Day 32
		09.00 h	12.00 h	09.00 h
Group				
1	8	TRH	GRF	Arginine
2	12	GRF	TRH	Arginine
3	12	TRH	GRF+TRH	GRF
4	12	TRH	GRF+TRH	Arginine
5	8	GRF+TRH	GRF	Arginine
6	8	GRF+TRH	TRH	Arginine
7	8	GRF+TRH	GRF	TRH
8	12	GRF	GRF+TRH	Arginine

GRF, human pancreatic GH-releasing factor fragment; TRH, thyrotrophin-releasing hormone.

for 24 h. Separation was by SAC-CEL second antibody (Immunodiagnostics Systems, Boldon, Tyne and Wear, U.K.). Tenfold dilution of plasma did not affect accuracy, when readings were within the range of the assay (0·2–50 ng/tube). Standards and tracer were made up from recombinantly derived bovine somatotrophin (Cyanamid, Princeton, NJ, U.S.A.), radioiodination was by the iodogen method (Pierce Chemicals, Poole, Dorset, U.K.). Quality control of the assay showed inter- and intra-assay variations of 12·9% ($n=9$) and 10·6% ($n=68$) respectively.

Serum from samples taken before the administration of the secretagogues were analysed for triglyceride concentration using a commercial kit (Roche Products Ltd, Welwyn Garden City, Herts, U.K.).

Statistics

Concentrations in single samples were combined into baseline concentration (mean of samples taken at −15 and −5 min), and into response (mean of 5-, 10-, 15- and 20-min samples) following GRF, TRH and the combined dose of GRF and TRH, or the mean of 10-, 15-, 20- and 30-min samples following arginine. Baseline and response variables were both transformed using natural logarithms in order to obtain errors that were approximately normally distributed. The response to each secretagogue was analysed separately for effects of experimental factors: i.e. sex of calf, diet and group. The effects of selection line and age (deviation from mean) were fitted as linear regressions, the former using the PBV of each calf. The effects of liveweight, baseline concentration and triglyceride concentration prior to administration on the response were investigated by covariance analysis (Urquhart, 1982). The effects of sires within selected lines were considered as random in the analysis, but they were not significant once PBV had been fitted.

RESULTS

Concentrations of GH before administration of any of the secretagogues did not differ between diets, selection lines, sexes or groups of calves, and averaged 8·82±0·43 (S.E.M.) μg/l. Following the administration of GRF, TRH, the combined dose of GRF+TRH and arginine, peak concentrations of GH were 34·7±4·2, 21·3±1·8, 180·1±24·1 and 18·9±1·9 μg/l respectively averaged across experimental factors. The response to GRF, TRH and GRF+TRH was immediate, with peaks occurring at 20, 10 and 10 min after administration respectively, whereas the onset of response to arginine was delayed 5–10 min and the peak did not occur until 30 min after (Figs 1 and 2). Following the peak, GH concentrations declined asymptotically towards preinjection levels within the 2-h period observed and secondary peaks were not detected.

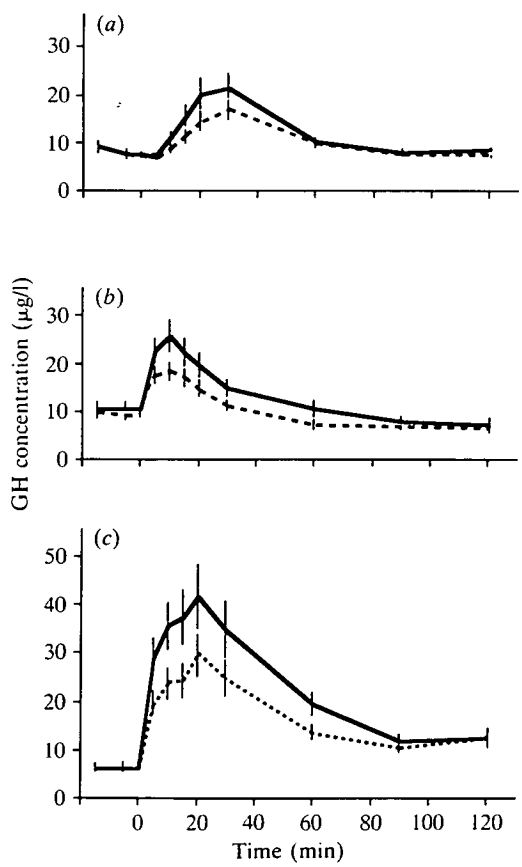


FIGURE 1. Concentration of GH before and after the administration at time 0 of (*a*) arginine, (*b*) TRH and (*c*) GH-releasing factor to calves of genetically high (solid lines) and low (broken lines) merit for milk yield. Values are means ± S.E.M.

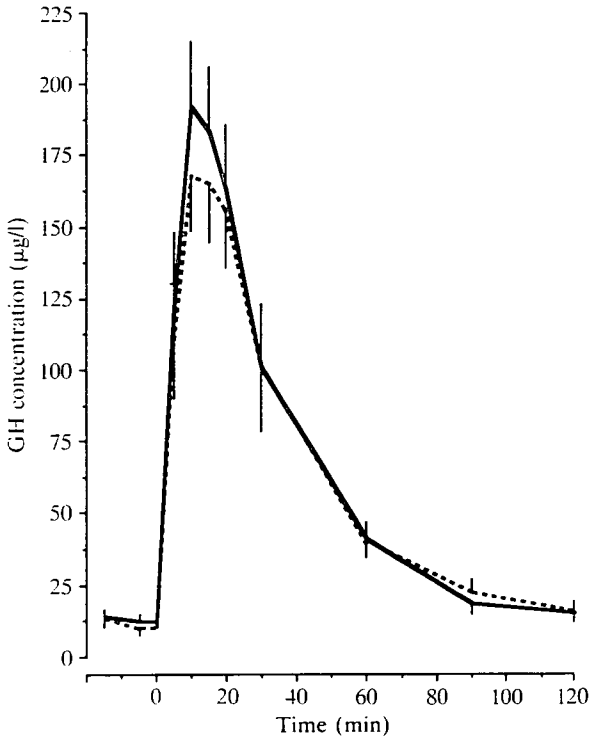


FIGURE 2. Concentration of GH before and after the simultaneous administration at time 0 of GH-releasing factor and TRH to calves of genetically high (solid lines) and low (broken lines) merit for milk yield. Values are means ± S.E.M.

The coefficients of the regressions of the GH response ($\log_e$ units) on PBV were $0{\cdot}0061 \pm 0{\cdot}0022$ ($P<0{\cdot}01$) for GRF, $0{\cdot}0042 \pm 0{\cdot}0020$ ($P<0{\cdot}05$) for TRH, $0{\cdot}0015 \pm 0{\cdot}0023$ ($P>0{\cdot}1$) for GRF and TRH combined and $0{\cdot}0033 \pm 0{\cdot}0018$ ($P<0{\cdot}1$) for arginine. Figures 1 and 2 show the mean line difference for each secretagogue at each sampling point.

The response was found to be highly correlated ($P<0{\cdot}001$) with the serum concentration before the injections. The log–log regressions of GH response on concentration prior to injection were $0{\cdot}43 \pm 0{\cdot}10$, $0{\cdot}54 \pm 0{\cdot}08$, $0{\cdot}33 \pm 0{\cdot}09$ and $0{\cdot}20 \pm 0{\cdot}08$ following GRF, TRH, GRF+TRH and arginine respectively. These represent, respectively, 1·35-fold, 1·45-fold, 1·30-fold and 1·15-fold increases in response for a doubling of the concentration prior to injection. The influence of concentration prior to injection did not account for the observed regressions of GH response on PBV since there was no association of PBV with concentration evident prior to injections.

Females responded 2·3-fold ($P<0{\cdot}001$) more than males in response to GRF alone; in contrast, however, females only responded 1·2-fold more than males when given TRH alone ($P>0{\cdot}1$; i.e. not significantly different from 1) and significantly ($P<0{\cdot}05$) less strongly when given GRF and TRH in combination (0·63-fold). Responses of the sexes to arginine were more complex, with the more responsive sex differing from group to group (interaction, $P<0{\cdot}01$). There was no difference between the sexes in the magnitude of the regression coefficient for GH response on PBV.

For the three treatments involving TRH and GRF there was no dietary effect observed ($P>0{\cdot}1$). However, the picture was more complex when arginine was used as a secretagogue ($P<0{\cdot}01$); in calves fed diets 2 and 3, GH responses were approximately 1·88-fold greater. The magnitude of the regression coefficient for GH response on PBV was not influenced by diet.

The grouping of calves was a necessary blocking factor, and different groups of calves were subject to differing ambient temperatures and light intensities through seasonal variation. Mean ambient temperature ranged between 4·5 and 12·5 °C and light intensity between 14 and 31 lx (both measured just before administration of secretagogues). Furthermore, all interassay variation was confounded with groups. Age, liveweight and triglyceride concentrations prior to injection were not significantly associated with GH responses. There was a non-significant trend ($P>0{\cdot}05$) for GH responses to be smaller when GRF or TRH was administered after their combination on day 30. However, there was no evidence that the magnitude of the regression coefficient for GH responses on PBV differed from group to group.

DISCUSSION

The results of this trial have provided evidence (i) for a positive association of the release of GH following GRF administration with the predicted breeding value for milk yield, (ii) for a similar effect following TRH administration, (iii) that the individual responses obtained depended critically on concentrations prior to secretagogue administration, and (iv) that at 4 months of age females respond more readily to GRF than males but not when it was given in combination with TRH.

This is the first report of a clear positive association of GH release following GRF with PBV and the largest study to observe such an effect of TRH. Previously Massri, Al-Raheem, Young & Wheaton (1985) carried out a smaller scale trial using Holstein heifers of differing genetic merit for milk yield. Although, from the figure of mean responses shown in their paper, it can be seen that heifers of high merit released more GH than heifers of low merit, the method of statistical analysis (choosing the maximum GH concentration observed for each animal after injection regardless of time) seemed to obscure this

result. Two smaller studies have reported similar findings with TRH (Barnes, Kazmer, Pearson & Akers, 1984; Løvendahl, Sejrsen & Andersen, 1989).

The doses of GRF(1–29) and TRH were chosen to give an intermediate degree of response. Petitclerc, Pelletier, Lapierre *et al.* (1987) showed from dose–response trials of both human GRF(1–29) and human GRF(1–44) with 4-month-old female calves that the fractions were equipotent in equimolar doses and that 0·067 nmol/kg liveweight (equivalent to 0·23 μg GRF(1–29)/kg liveweight) was approaching the region of the curve with diminishing return in response. Johke, Hodate, Ohashi *et al.* (1984) showed in females of a similar age that most of the response to GRF(1–44) occurred between 0·1 and 0·5 μg/kg liveweight (equimolar with 0·07 and 0·33 μg GRF(1–29)/kg liveweight). These indicate 0·2 μg GRF(1–29) is close to the half-maximum effective dose, an observation consistent with other studies of dose–response in cattle of different ages. The results of Hedlund, Doelger, Tollerton *et al.* (1977) in 5-month-old male calves suggest that maximal response to TRH is much less than to GRF and that beyond 0·25 μg/kg liveweight response is relatively insensitive to dose. The size of the response may explain the poorer relationship of GH release with PBV following the combination of GRF and TRH. A critical experiment would therefore involve the use of more than one dose of GRF or TRH spanning the dose–response curves to separate differential sensitivity and/or maximal response.

It has been a consistent finding of studies both within and between breeds that genetically high-yielding cows have greater circulating concentrations of GH during lactation (Hart *et al.* 1978; Flux, Mackenzie & Wilson, 1984; Kazmer, Barnes, Akers & Pearson, 1986; Bonczek, Young, Wheaton & Miller, 1988; Lukes, Barnes & Pearson, 1989). Kazmer *et al.* (1986) showed that the release of GH following TRH administration was greater in lactating cows of high genetic merit for milk yield and Lukes *et al.* (1989) showed that this was also the case with GRF. Whilst these results indicate that GH plays a role in mediating genetic improvement in lactation yields, there remains, however, the key question as to whether this arises indirectly through the increase in net energy deficit that is also observed. The results of this trial therefore give a strong indication that the differences in GH observed are a direct result of selection pressure.

The strong differences in the nature of sexual dimorphism with GRF and GRF+TRH emphasize the differences in mechanism of the secretagogues. There is evidence in rats to suggest that the response to GRF is greater in females before puberty (Heiman, Nekila, Murphy *et al.* 1984; Lopez, Gonzales & Aguilar, 1986) and in males following puberty (Evans, King, Limber *et al.* 1985), with testosterone stimulating release (Wehrenberg, Baird, Ying & Ling, 1985). This trial shows that at least in prepubertal calves, females can respond to GRF more than males; a result in agreement with evidence from the rat. The opposite sex effect observed with the synergistic response of TRH and GRF combined is a new result and in the absence of detailed knowledge of the mechanism involved cannot be deduced from previous results.

In conclusion, the results have provided evidence that the positive association observed in dairy cows between circulating GH concentration and genetic merit for milk yield during lactation has its counterpart in calves in the release of GH following administration of GRF, and to a lesser extent TRH.

ACKNOWLEDGEMENTS

This work was supported by the U.K. Ministry of Agriculture, Fisheries and Food, the Underwood Fund of the Agricultural and Food Research Council and the Danish Research Academy whose grant made the visit of P.L. to the Institute of Animal Physiology and Genetics Research possible. The authors are grateful to Mr B. Wilson for all his skilled work in all aspects of obtaining the blood samples, Mr S. Nisbet for his help in the collection and analysis of samples, Dr J. Oldham of the East of Scotland College of Agriculture for his advice on the formulation of the diets, and to the staff of Dryden-Mountmarle and Blythbank farms for their care of the calves.

REFERENCES

Agricultural Research Council (1980). *The Nutrient Requirements of Ruminant Livestock*. Slough: Commonwealth Agricultural Bureaux.

Anfinson, M. S., Davis, S. L., Christian, E. & Everson, D. O. (1975). Episodic secretion of growth hormone in steers and bulls: An analysis of frequency and magnitude of secretory spikes occurring in a 24 hour period. *Proceedings of the Western Section, American Society of Animal Science* **26**, 175–177.

Barnes, M. A., Kazmer, G. W., Pearson, R. E. & Akers, R. M. (1984). Endocrine and metabolite response after insulin, glucose or TRH in two genetically selected groups of Holstein cattle. *Journal of Dairy Science* **67** (Suppl. 1), p. 161, Abstract No. 156.

Bonczek, R. R., Young, C. W., Wheaton, J. E. & Miller, K. P. (1988). Responses of somatotropin, insulin, prolactin and thyroxine to selection for milk yield in Holsteins. *Journal of Dairy Science* **71**, 2470–2479.

Chilliard, Y. (1989). Long term effects of recombinant bovine somatotropin (rBST) on dairy cow performances: A review. In *Use of Somatotropin in Livestock Production*, pp. 61–87. Eds K. Sejrsen, M. Vestergaard & A. Neimann-Sorensen. Amsterdam: Elsevier.

Evans, W. S., King, R. J., Limber, E. R., Kaiser, D. L. & Thorner, M. O. (1985). Effects of *in vivo* gonadal hormone environment on

in vitro hGRH-40-stimulated GH release. *American Journal of Physiology* **249**, E276–E280.

Flux, D. S., Mackenzie, D. D. S. & Wilson, G. F. (1984). Plasma metabolite and hormone concentrations in Friesian cows of differing genetic merit measured at two feeding levels. *Animal Production* **38**, 377–384.

Hart, I. C., Bines, J. A., Morant, S. V. & Ridley, J. L. (1978). Endocrine control of energy metabolism in the cow: comparison of the levels of hormones (prolactin, growth hormone, insulin and thyroxine) and metabolites in the plasma of high- and low-yielding cattle at various stages of lactation. *Journal of Endocrinology* **77**, 333–345.

Hart, I. C., Flux, D. S., Andrews, P. & McNeilly, A. S. (1975). Radioimmunoassay for ovine and caprine growth hormone: its application to the measurement of basal circulating levels of growth hormone in the goat. *Hormone and Metabolic Research* **7**, 35–40.

Hedlund, L., Doelger, S. G., Tollerton, A. J., Lischko, M. M. & Johnson, H. D. (1977). Plasma growth hormone concentrations after cerebroventricular and jugular injection of thyrotropin-releasing hormone. *Proceedings of the Society for Experimental Biology and Medicine* **156**, 422–425.

Heiman, M. L., Nekila, M. V., Murphy, W. A., Lance, V. A. & Coy, D. H. (1984). Development of prepubertal pituitary responsiveness to growth hormone-releasing factor and luteinizing hormone-releasing hormone. *Biology of Reproduction* **30** (Suppl. 1), p. 180, Abstract No. 299.

Johke, T., Hodate, K., Ohashi, S., Shiraki, M. & Sawano, S. (1984). Growth hormone response to human pancreatic growth hormone releasing factor in cattle. *Endocrinologia Japonica* **31**, 55–61.

Kazmer, G. W., Barnes, M. A., Akers, R. M. & Pearson, R. E. (1986). Effect of genetic selection for milk yield and increased milking frequency on plasma growth hormone and prolactin concentration in Holstein cows. *Journal of Animal Science* **63**, 1220–1227.

Lopez, F., Gonzalez, M. & Aguilar, E. (1986). GH response to GRF 1–29: effects of age, sex, cycle phase and testosterone pretreatment. *Journal of Steroid Biochemistry* **25**, p. 81S, Abstract No. 216.

Løvendahl, P., Sejrsen, K. & Andersen, B. B. (1989). Release of growth hormone as juvenile predictor of dairy merit. In *New Selection Schemes in Cattle: Nucleus Programmes*, pp. 134–139. Wageningen: Pudoc.

Lukes, A. J., Barnes, M. A. & Pearson, R. E. (1989). Response to selection for milk yield and metabolic challenges in primiparous dairy cows. *Domestic Animal Endocrinology* **6**, 287–298.

McAtee, J. W. & Trenkle, A. (1971). Effect of feeding, fasting and infusion of energy substrates on plasma growth hormone levels in cattle. *Journal of Animal Science* **33**, 612–616.

Massri, Y. G., Al-Raheem, S. N., Young, C. W. & Wheaton, J. E. (1985). Growth hormone response to growth hormone-releasing factor in selected and control Holstein heifer calves. *Journal of Animal Science* **68**, 1819–1821.

Peel, C. J. & Bauman, D. E. (1987). Somatotropin and lactation. *Journal of Dairy Science* **70**, 474–486.

Petitclerc, D., Pelletier, G., Lapierre, H., Gaudreau, P., Couture, Y., Dubreuil, P., Morisset, J. & Brazeau, P. (1987). Dose response to two synthetic human growth hormone-releasing factors on growth hormone release in heifers and pigs. *Journal of Animal Science* **65**, 996–1005.

Takano, K., Hizuka, N., Shizume, K., Asakawa, K., Miyakawa, M., Hirose, N., Shibasaki, T. & Ling, N. C. (1984). Plasma growth hormone (GH) response to GH-releasing factor in normal children with short stature and patients with pituitary dwarfism. *Journal of Clinical Endocrinology and Metabolism* **58**, 236–241.

Urquhart, N. S. (1982). Adjustment in covariance when one factor affects the covariate. *Biometrics* **38**, 651–660.

Wehrenberg, W. B., Baird, A., Ying, S. & Ling, N. (1985). The effects of testosterone and estrogen on the pituitary growth hormone response to growth hormone-releasing factor. *Biology of Reproduction* **32**, 369–375.

Woolliams, J. A. & Smith, C. (1988). The value of indicator traits in the improvement of dairy cattle. *Animal Production* **46**, 333–345.

Paper 7

Anim. Prod. 1993, 56: 1-8

0003-3561/93/20120001$02.00

Endogenous pulsing and stimulated release of growth hormone in dairy calves of high and low genetic merit

J. A. Woolliams, K. D. Angus and S. B. Wilson

Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

Abstract

Fifty-five calves aged 105 days of age, of both sexes, belonging to two genetic groups (H, high or L, low) differing in their predicted breeding value (PBV) by 78 kg for fat plus protein yield, were individually penned for 3 weeks with the final 2 weeks on a diet designed to provide energy close to their maintenance requirements. The calves were then cannulated in the jugular vein and blood was sampled every 15 min for 25 h. Following this a growth hormone releasing factor (GRF) preparation was administered on up to four occasions, being one of (per kg live weight) either 0·2 or 0·4 μg GRF (treatments SGRF or DGRF respectively) or 0·2 μg thyrotropin releasing hormone (TRH). On each occasion blood samples were taken at −15, 4, 8, 12, 16, 20 and 45 min relative to the time of administration and up to four further occasions, one prior to and three within 32 min of administration. Samples were assayed for growth hormone (GH).

GH concentration of troughs prior to an episode of GH release was 1·19-fold greater in H compared with L calves with concentrations decreasing to 0·9 of their previous values each h. Trough and peak concentrations had repeatabilities of 0·21 and 0·26 respectively (both $P < 0·05$). There was no association between PBV and either peak concentrations, number of pulses (4·84 per 25 h) or mean GH concentration (13·3 μg/l).

GH released was only 1·11-fold greater after DGRF than SGRF. The difference in the regression coefficients for PBV between SGRF and DGRF was small, as it was for the difference between coefficients for the two sexes. The pooled coefficient was 0·00342 $\log_e$ units per kg (s.e. 0·00157; $P < 0·05$) indicating a 1·31-fold greater response in H than in L calves. The correlation between responses to SGRF and DGRF was 0·482 and the repeatability of SGRF was 0·338; a pooled repeatability was estimated as 0·362 ($P < 0·01$). The regression of GH release on PBV for TRH was 0·00345 (s.e. 0·00330). For all secretagogues, response depended on prior concentrations.

It was concluded that GH release following GRF administration was positively related to PBV in dairy calves and response was moderately repeatable; furthermore, although aspects of endogenous secretion may be related to PBV they suffer from measurement difficulties.

Keywords: *dairy calves, genetic merit, somatotropin.*

Introduction

Løvendahl, Angus and Woolliams (1991a) showed that, in dairy calves, the release of growth hormone (GH) following administration of growth hormone releasing factor (GRF) or thyrotropin releasing hormone (TRH) may be related to their breeding value for fat plus protein yield. Their study had been prompted by the observation of greater circulating GH (Kazmer, Barnes, Akers and Pearson, 1986; Bonczek, Young, Wheaton and Miller, 1988) and stimulated release of GH (Kazmer *et al.*, 1986) in lactating cows differing in genetic merit.

In their study a clear association was evident between the size of GH release and the concentration of GH prior to the administration of the secretagogue. This indicated that release may be influenced by the characteristics of the endogenous pulsing of GH, and if so a greater understanding of the result might be obtained by examining this relationship more closely. No published study has examined GH pulsing in calves of genetically high and low merit.

One application of differential GH release would be to incorporate this finding into breeding schemes as

an early predictor of merit in both sexes (Woolliams and Smith, 1988). For this use greater information is required in the genetic correlation of GH release with genetic merit for fat plus protein yield, and further information on the value of repeat testing as measured by the repeatability.

Therefore this experiment was designed with the following objectives: (i) to characterize fully the pulsatile release of endogenous GH in juveniles of high and low breeding value for fat plus protein yield; (ii) to confirm the association of breeding value with GH secretion following administration of secretagogues; (iii) to estimate the repeatability of GH secretion following administration of secretagogues; (iv) to obtain preliminary information on the effect of GRF dose on the association with breeding value; (v) to use information on the timing of pulses to understand more fully the association of GH responses to secretagogues with GH concentrations prior to administration.

Material and methods

The experiment was conducted using an initial group of seven calves and six further groups each of eight calves. Groups 1 and 2 were studied using one protocol whilst the remaining groups were studied using a revised protocol (see Table 1).

Fifty-five British Friesian calves, 32 females and 24 males, born over a period of 12 months were obtained from Blythbank Farm, Tweeddale and Cold Norton Farm, Staffordshire, and all were maintained at the former from 1 week of age. Calves were weaned by 12 weeks of age. All calves were sired by either one of nine high (H) merit or one of nine low (L) merit bulls. The H panel contributed 28 calves with a range of one to six calves per bull, whilst the L panel contributed 27 calves with a range of two to six calves per bull. Genetic information on maternal sires was known for a minimum of two and a maximum of six generations. The mean difference in predicted breeding age value (PBV) for the yield of fat plus protein of calves from H and L lines was 78 kg with a difference of 27·5 kg between the minimum of the H line and the maximum of the L line.

Table 1 *The allocation of growth hormone releasing factor (GRF) treatments according to groups (SGRF and DGRF represent 0·2 and 0·4 μg GRF per kg live weight respectively and TRH represents 0·2 μg TRH per kg live weight)*

	Day 3		Day 4	
	09.30 h (A)	14.30 h (B)	09.30 h (C)	14.30 h (D)
Group				
1	SGRF		SGRF	TRH
2	SGRF		SGRF	TRH
3	SGRF	SGRF	SGRF	DGRF
4	SGRF	DGRF	SGRF	SGRF
5	SGRF	DGRF	SGRF	SGRF
6	SGRF	SGRF	SGRF	DGRF
7	SGRF	DGRF	SGRF	SGRF

In groups of eight at an average of 105 days of age (s.d. 9·8, range 84 to 129) the calves were individually penned at Dryden Farm, Midlothian. In each group every calf of the H line was penned adjacent to a calf of the L line of the same sex and similar age. The subsequent biochemical analysis always included such pairs in the same assay. The calves were offered 2 kg concentrate and 1·7 kg hay per day. The concentrate was a commercial calf rearing diet (Langhill Growers; East of Scotland College of Agriculture) and was given in two equal portions in the morning and afternoon. Seven days later the concentrate portion of the ration was reduced to a total of 0·64 kg concentrate again given in two equal portions. This ration was designed to provide the average maintenance requirement for energy of males and females weighing 120 kg (Agricultural Research Council, 1980). It was maintained until the calves had completed the experimental schedule. Fourteen days after the introduction of the maintenance ration the calves were cannulated (day 1) in the jugular vein. At the start of the period on reduced rations, mean live weight was 111·0 kg and on the day of cannulation had increased to 116·8 kg, with no significant difference between the lines (H-L, 4·9 kg; s.e. 2·9).

Starting from 08.00 h on day 2 the calves were bled every 15 min for a period of 25 h, i.e. up to and including 09.00 h on day 3. Calves were then administered a GRF preparation on up to four occasions: day 3, 09.30 h (A); day 3, 14.30 h (B); day 4, 09.30 h (C); and day 4, 14.30 h (D). Table 1 shows the timing of administration for each group. Calves in groups 1 and 2 (i.e. on the original protocol) did not receive a GRF preparation on occasion B. On each of these occasions samples were always taken (relative to time of administration) at −15, 4, 8, 12, 16, 20 and 45 min; and also for groups 1 and 2 at 30 and 60 min, but for groups 3 to 7 the sequence included samples at −5, 24, 28 and 32 min. For all groups, 0·2 μg GRF [(1-29)NH_2] per kg live weight (treatment SGRF) was administered on occasions A and C. For groups 1 and 2, 0·2 μg TRH per kg live weight was administered on occasion D (Table 1). These doses of GRF and TRH were those used by Løvendahl *et al.* (1991a). GRF was obtained from Bachem (Saffron Walden) and TRH from Sigma (Poole, Dorset).

The revision of the treatment protocol after the first two groups made the following changes: the use of TRH was discontinued; SGRF was administered to each calf three times; a dose of 0·4 μg GRF per kg (treatment DGRF) was administered to each calf once either on occasion B or D (Table 1). The choice was made at random for each group.

For group 7, within 8 h of the start of the serial sampling on day 2, six of the eight cannulae had stopped working. It was decided to re-cannulate all eight calves which was completed by 17.00 h on day 2. Serial sampling did not resume until 03.00 h on day 3.

Analytical methods

The samples were assayed for GH using the method described by Løvendahl, Woolliams and Sinnett-Smith (1992b). The GH pulse sequences for up to four calves were analysed in each assay and these were made up of two pairs from a single group. All of the samples taken during the challenges for a single group were assayed together irrespective of the preparation or dose given. The quality control samples of high concentration (16·9 μg/l) in the assays had inter- and intra-assay coefficients of variation of 0·119 (no. = 22) and 0·099 (no. = 208) respectively, and the corresponding values for the quality control of low concentration (11·0 μg/l) were 0·140 (no. = 22) and 0·114 (no. = 212).

Statistical methods

The analysis of the pulses using parametric methods (Diggle and Zeger, 1989) was investigated. However, it was found these models were not robust to the shape of the pulses in this trial; the pulses consisted not only of classical high concentration pulses followed by an exponential decay, but also more sustained episodes of GH release.

The analysis, therefore, used a non-parametric approach (modified from Breier, Bass, Butler and Gluckman, 1986). In this approach the analysis defined a *shift* in a sequence of GH concentration to have occurred when the second of two samples exceeded the first by at least α units on a logarithmic scale. The value of α used was $\chi_{0.99}\sigma$ where σ was the standard deviation of the difference between the high and low quality controls (log values) measured within assays and $\chi_{0.99}$ is the upper 1% truncation point of a N(0, 1) distribution. α therefore serves as a measure of extraneous variation.

It was considered necessary to define further the occurrence of a pulse to protect against (i) single rogue samples and (ii) identifying a sustained rise over several samples as more than a single pulse. Therefore the ith pulse was defined to have occurred at sample n when: (a) a *shift* was present between samples $(n-2)$ and n, $(n-1)$ and n, $(n-2)$ and $(n+1)$, and, $(n-1)$ and $(n+1)$; (b) no *shift* had been identified at times $(n-3)$, $(n-2)$ and $(n-1)$; and (c) the mean of samples $(n-2)$ and $(n-1)$ was less than the mean of the two upper samples of the previous pulse.

For the ith pulse of a sequence the following were defined: (a) its time (n_i); (b) the mean of samples (n_i-2) and (n_i-1), the step trough; (c) the mean of samples n_i and (n_i+1), the step peak; (d) for $i>1$ the minimum concentration of a single and a consecutive pair of samples between pulses $(i-1)$ and i and times of occurrence; (e) the maximum concentration of a single and consecutive pair of samples between the pulses i and $(i+1)$ and the times of occurrence. For the entire sequence of an individual the following were defined for subsequent use; the mean GH concentration, the number of pulses, the time of the last pulse in the sequence and the time and magnitude of the subsequent peak concentration.

The parameters of each pulse were analysed using residual maximum likelihood (Patterson and Thompson, 1971) with a mixed linear model that incorporated fixed effects for group (5 d.f., since group 7 was excluded) and sex (1 d.f.), with age and PBV as covariates. The random components of the model assumed an error term of the form:

$$v_{ijk} = \alpha_i + \beta_{ij} + \theta_{ijk}$$

for the kth pulse of the jth offspring of the ith sire and where α_i, β_{ij} *and* θ_{ijk} are independent random variables with means of zero. After initial analysis the time from the previous peak was included as an additional covariate. The trough of the first pulse and peak of the final pulse were excluded from analyses since, strictly, these were undefined.

The mean concentration over the entire sequence, each individual sample and the number of pulses observed in each sequence were analysed using linear models with the same fixed effects as the previous analysis. The analysis of the grand mean and samples on each occasion required only a basic random model of homogeneous uncorrelated normal errors but for the number of pulses the linear model was fitted to the logarithm of the number observed and Poisson errors were assumed on the untransformed scale.

The results from the challenges with the releasing factor preparations were combined as described by Løvendahl *et al.* (1991a). The mean of the samples at −15 and −5 min was taken as an estimate of GH concentration prior to injection, and the mean of samples at 4, 8, 12, 16 and 20 min were taken as an

estimate of challenge reponse, although other measures were also investigated. The analysis of SGRF and DGRF responses were combined as far as possible to enable a close comparison.

Firstly, the baseline concentrations were analysed using residual maximum likelihood. A fixed effect model was fitted that accounted for batch (6 d.f.), sex (1 d.f.), time of challenge (i.e. A, B, C or D; 3 d.f.), dose (1 d.f.) and a sex × dose interaction (1 d.f.). Covariate adjustment was also made for age (1 d.f.) and PBV (1 d.f.) and, following an initial inspection of the data, for the interaction of time of challenge with PBV (3 d.f.). Since doses were yet to be applied a simple random model was fitted to error terms (v_{ijk}):

$$v_{ijk} = \alpha_i + \beta_{ij} + \theta_{ijk}$$

for the kth challenge of the jth offspring of the ith sire and where α_i, β_{ij} and θ_{ijk} are independent random variables with means of zero. The repeatability of baseline sampling was estimated as $t_b = (E(\alpha_i^2) + E(\beta_{ij}^2))/E(v_{ijk}^2)$. From these analyses v_{ijk} was subsequently used as a covariate in the analysis of mean responses.

The mean response to the challenge was also investigated using residual maximum likelihood. An initial analysis was made with no adjustment for baseline to investigate the error structure. This analysis fitted the same model as that fitted to baseline except the interaction of PBV with dose and sex was investigated, rather than with the time of challenge. The error terms were modelled by:

$$\varepsilon_{sij} = \gamma_i + \delta_{ij}$$

for the jth challenge of SGRF to the ith calf and by ε_{di} for the challenge of DGRF to the ith calf, where $E(\gamma_i) = E(\delta_{ij}) = E(\gamma_i\ \delta_{ij}) = 0$. The following variances and covariances were estimable: $E(\gamma_i^2)$, $E(\varepsilon_{sij}^2)$ and $E(\gamma_i\ \varepsilon_{ij})$. Due to the complexity of fitting this model, sire terms were not included.

From these repeatability of SGRF (t_s) was calculated and the correlation of responses to SGRF and DGRF (r):

$$t_s = E(\gamma_i^2)/E(\varepsilon_{sij}^2) \text{ and } r = E(\gamma_i\ \varepsilon_{di})/(E(\varepsilon_{sij}^2)\ E(\varepsilon_{di}^2))^{1/2}.$$

Following this analysis, regression on v_{ijk} from the analysis of baseline concentrations were added to the fixed model and interactions of PBV with time of challenge were also investigated.

Since the first challenge with releasing factor (occasion A) followed on immediately from the pulse bleed, further analyses were carried out to correlate response to secretagogue with endogenous pulsing. These analyses included various covariates derived from the analysis of pulses carried out previously. Covariates considered were the time and peak concentrations of the last pulse identified prior to the challenge and the mean concentration of all samples taken during the serial sampling. The simpler error structure for these analyses enabled analysis of variance to be used with fixed models that included effects for batch and sex as well as linear regressions PBV, age and the covariates described above.

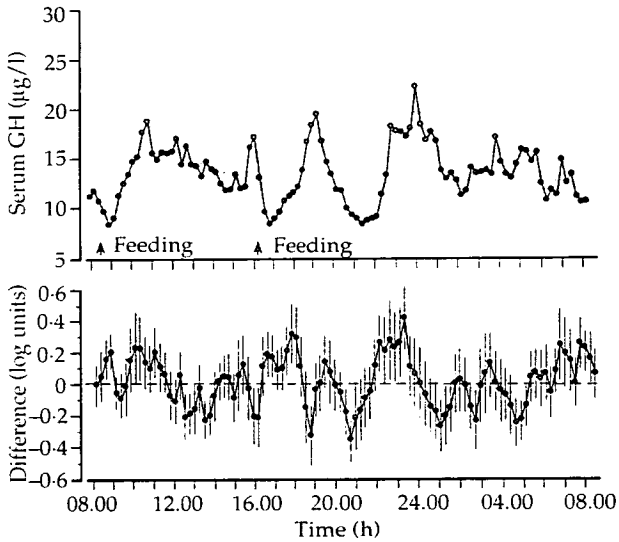


Figure 1 The mean GH concentration during 25 h of serial sampling, derived from all calves (no. = 47) with complete sequences (↑ indicates times of feeding) and the difference between calves of genetically high and low merit for milk yield. Stippled areas denote ± 1 s.e. on a log scale.

Results

Analysis of pulses

The average GH concentration over all calves with complete records at each sampling occasion is shown in Figure 1. It is clear that the timing of pulses is not random and that some common stimuli are present. Troughs are particularly clear just after feeding in both morning and afternoon and at 09.30 h and 01.30 h. There was no significant difference between H and L calves at any one occasion. Figure 2 is a histogram of all the pulse times that were identified pooled over the same individuals. The analysis of the number of pulses observed and the geometric mean GH concentration of the entire sequence also showed no difference between H and L calves (means of 4·84 pulses per 25 h, and 13·34 μg/l).

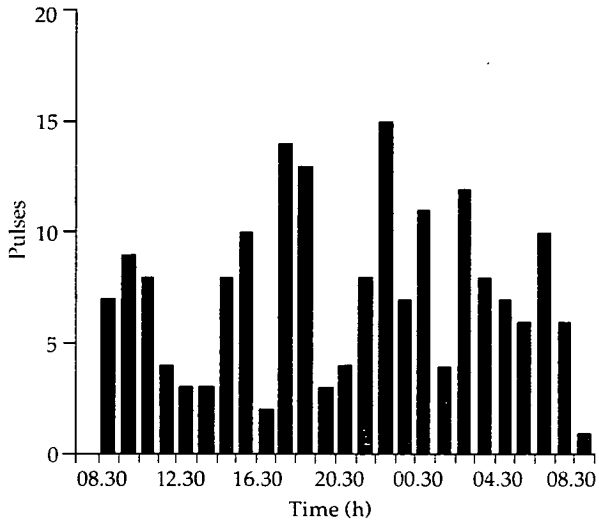


Figure 2 A histogram of the number of pulses observed in hourly periods for all calves (no. = 47) with complete sequences. An artefact of the technique of detection is that few pulses will be detected in the first group (i.e. prior to 08.30 h) and the last group.

Analysis of pulse parameters showed peak concentrations to have repeatabilities of 0·252 ($P < 0·05$), 0·258 ($P < 0·05$) and 0·175 ($P > 0·05$) for one-sample, two sample and step peaks respectively. No systematic variation with experimental factors was identified and peak concentrations were not affected by the interval from the previous pulse. Measures of trough concentrations were also repeatable; 0·235 ($P < 0·05$), 0·210 ($P < 0·05$) and 0·180 ($P > 0·05$) for one-sample, two-sample and step troughs respectively. In contrast to peaks, troughs were both positively associated with PBV ($P < 0·05$ for all three measures), and negatively associated with the time since the previous pulse ($P < 0·001$). Trough concentrations were 1·19-fold greater in H compared with L calves and decreased approximately 0·9-fold per h. The time since the previous pulse was not associated with PBV (as expected from the analysis of pulse frequencies) and the covariance adjustment of time increased the estimates of repeatability. As a result the regression coefficients for trough concentration on PBV and their standard errors were largely unaltered by adjustment. The different measures of peak and trough concentrations were highly correlated and consequently the results for different measures were very similar.

Analysis of challenges

The model fitted to baseline concentrations estimated the repeatability, $t_b = 0·247$ ($P < 0·05$). There was no effect of dose, or sex, or their interaction on baseline concentrations. GH concentrations in the afternoon were 1·43-fold greater than in the morning ($P < 0·001$). Baseline concentrations were positively associated with PBV and H calves had baseline concentrations 1·18-fold greater than those of L calves; the regression coefficient was 0·00208 $\log_e$ units per kg (s.e. 0·00095; $P < 0·05$). Although none of the coefficients describing the interaction of PBV with the time of the challenge was significant, the regression coefficients for each occasion ranged from 0·00044 to 0·00500 for occasions C and D respectively.

Analysis of the error structure estimated the total phenotypic variances as 0·4586 and 0·4654 for SGRF and DGRF respectively. The repeatability for SGRF was estimated as 0·338 ($P < 0·01$), but the correlation between responses to SGRF and DGRF was estimated as 0·482; i.e. greater than the correlation between two doses of SGRF. Thus subsequent analyses assumed $t_s = r$, and equal phenotypic variances. Under these assumptions, and after regression on baseline concentrations, the phenotypic variance (σ_c^2) was estimated as 0·399 and $t_s = 0·362$ ($P < 0·01$). The simplification of the random model enabled sire terms to be fitted but these proved to be small (the estimated variance component was not positive) and were consequently neglected.

Analysis of the mean response showed only a 1·11-fold greater response to DGRF than SGRF ($P > 0·05$). For both SGRF and DGRF the mean response was positively related to PBV and the difference between regression coefficients was small. Likewise, the

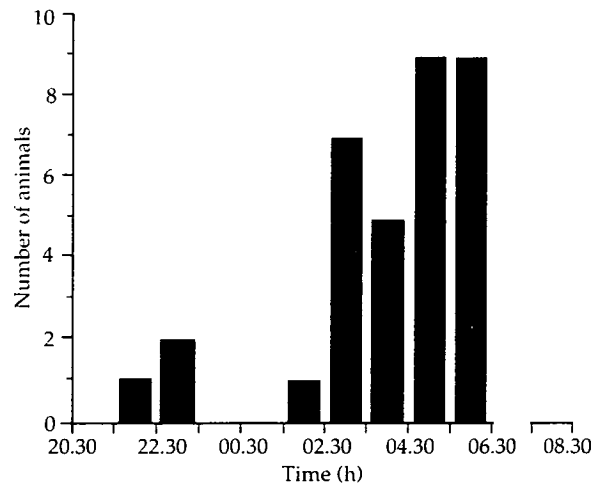


Figure 3 A histogram of the times when the last pulse was observed prior to GRF administration (09.30 h) on occasion A for all calves with complete records (no. = 47).

difference in the regression coefficient for response on PBV was 0·00342 $\log_e$ units per kg (s.e. 0·00157; $P < 0.05$) indicating a 1·31-fold greater response in H calves compared with their L counterparts. There was no significant interaction of regression on PBV with time of challenge and the coefficients varied from 0·00182 to 0·00710 for occasions C and D respectively.

The regression on baseline error was 0·589 (s.e. 0·122). The magnitude indicates that a doubling of the baseline concentration would be associated with a 1·50-fold increase in response. When analysis was restricted to the first challenge with GRF at occasion A there was no evidence for either a linear or quadratic association with either the time or magnitude of the preceding pulse. The distribution of the intervals between the preceding pulse and challenge are given in Figure 3. The mean interval was 200 min. The regression coefficient of response at occasion A on the grand mean of the serial sampling was 0·89 (s.e. 0·60). When both the grand mean and deviations in baseline unaccounted for by differences in grand mean (calculated by regression analysis) were included as covariates the regression coefficients were 0·89 (s.e. 0·60) and 0·69 (s.e. 0·30) respectively. The regression coefficient of baseline concentration on the individual grand means was 0·91 (s.e. 0·33).

The shape of the response for SGRF, DGRF and TRH is shown in Figure 4. On average SGRF reached peak concentrations 16 min after administration and had substantially decreased by 45 min after administration. DGRF achieved peak concentration 12 min after administration. The difference between SGRF and DGRF remained small throughout.

The response for TRH was clearly smaller and the regression of response on PBV was 0·00345 $\log_e$ units per kg (s.e. 0·00330). The large s.e. was due to the small number of observations.

Discussion

The first objective was a more complete description of endogenous GH profiles in H and L calves. This trial has provided evidence that the GH concentration in troughs prior to an episode of GH release (but not the following peak concentration) was related to predicted breeding value for fat plus protein production. The absence of an association with peak concentration must be treated with caution since the serial sampling procedure is effective in identifying the occurrence of a pulse and estimating the slowly varying trough concentration but is relatively poor in identifying the peak during a

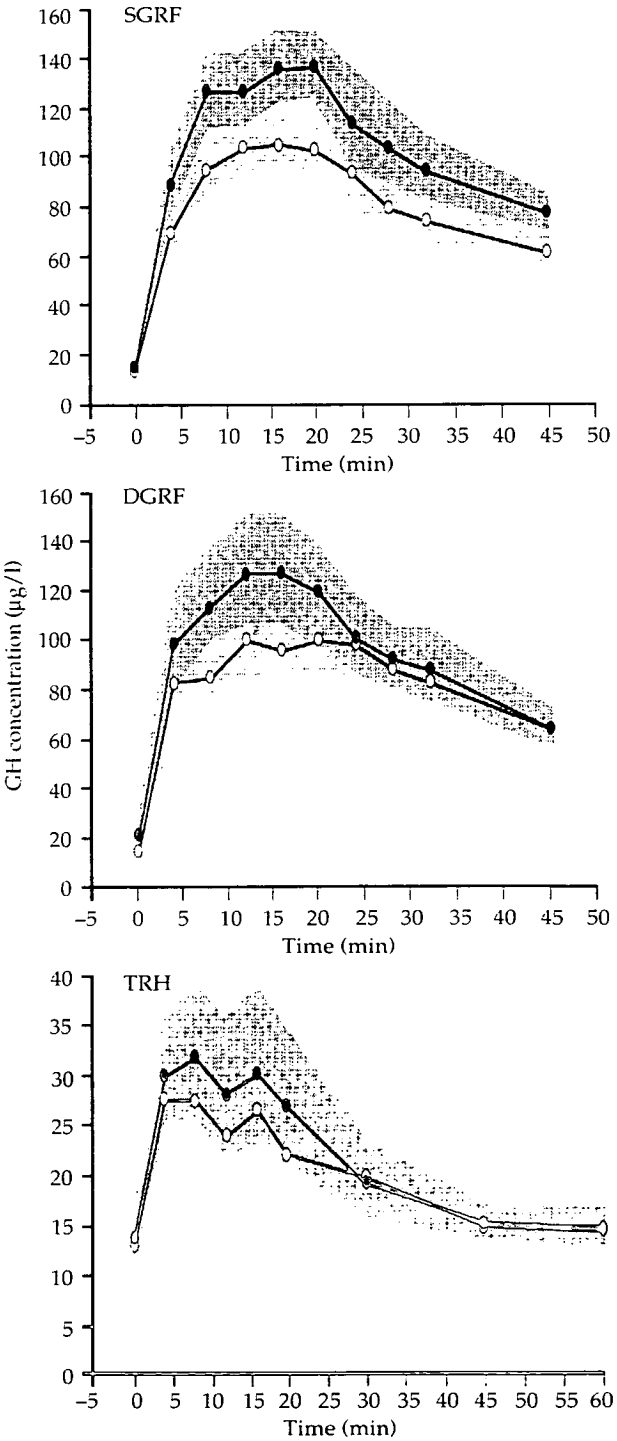


Figure 4 Concentration of GH before and after administration of secretagogues at time 0 of calves of genetically high (●) and low (○) merit for milk yield. Stippled areas denote ± 1 s.e. on a log scale.

period of rapid change. Nevertheless, the observation of an association with concentrations during troughs coupled with the observation that a degree of synchrony can be apparent in the occurrence of GH pulses, may be sufficient to explain why some studies (Barnes, Kazmer, Akers and Pearson, 1985; Mackenzie, Wilson, McCutcheon and Peterson, 1988) have observed differences between lines of cattle selected for milk yield without full-scale sequential sampling. Indeed, such a difference was observed prior to the challenges. However, there was no difference between lines in the mean GH concentration over 25 h, the number of pulses observed in the 25 h or at any single occasion during the 25 h. This mitigates against the use of endogenous GH concentrations as a predictor unless reliable environmental stimuli of episodic release can be identified and controlled. Some harmonizing stimuli have been suggested : light (Tannenbaum and Martin, 1976) and feeding (Mosely, Alaniz, Clafin and Krabill, 1988) although both the trials mentioned had the stimuli confounded with time of day. This trial was not designed to identify such stimuli.

Two objectives achieved by the trial were the confirmation of the positive association of PBV with GH release following administration of GRF and the obtaining of preliminary information on the effect of the dose of GRF on this association. Doubling the dose of GRF gave only a 1·11-fold increase in response; however, the correlation of the responses of an individual to SGRF and DGRF was observed to be as great as that for two challenges with SGRF, and the regression coefficients of response on PBV were almost identical. The pooled regression coefficient of 0·00342 (s.e. 0·00157) was lower than that observed by Løvendahl *et al.* (1991a) although the associated s.e.s. indicate the estimates are compatible. However, there are differences in experimental protocols which might have contributed to this result: in this trial the calves were given maintenance rations for 2 weeks prior to the challenges compared with normal feeding for 4 days after a 72-h fast. This decision was deliberate in order to tackle more effectively the characterization of endogenous pulses. Breier *et al.* (1986) have shown in steers, and Thomas, Mercer, Karalis, Rao, Cummings and Clarke (1990) in adult sheep, a more marked pulsatility on restricted feeding with greater pulse amplitude, and this same phenomenon may be responsible for the greater release of GH following administration of SGRF in this trial compared with that of Løvendahl *et al.* (1991a). Over the two experiments, there is an empirical, convex relationship between the magnitude of the regression on PBV and the size of response (Figure 5). Re-analysis of selected samples from this trial and that of Løvendahl *et al.* (1991a) excluded the possibility that part of the difference in magnitude of response to SGRF between them may be due to inter-assay variation.

A further objective was to estimate the repeatability of response to SGRF and this was estimated after fixed model adjustment (including baseline) as 0·328 with a phenotypic variance of 0·374 (c.f. 0·358 for P. Løvendahl (1991), unpublished results). Taken together, the results of this experiment and that of Løvendahl *et al.* (1991a) enable estimates of genetic parameters for the relationship of GH-release with breeding value for fat and protein production to be made. Firstly, assume the regressions on PBV in the two trials are estimates of the same coefficient. The pooled value is 0·00440 (s.e. 0·00131) and estimates $(r_G h_C \sigma_C)/(h_M \sigma_M)$ where σ_C^2 and σ_M^2, h_C^2 and h_M^2 are the phenotypic variances and heritabilities for the challenge and fat plus protein yield respectively, and r_G is the additive genetic correlation. Assume h_M^2 = 0·25, σ_M = 70 kg and σ_C = 0·60 log units. Consequently $h_C r_G \approx 0{\cdot}26$ and $h_C r_G h_M \approx 0{\cdot}13$. From the present trial $h_C^2 < t_s$ = 0·328 and thus $r_G > 0{\cdot}4$. However, there is a wide confidence interval on the estimate of the regression coefficient and even the s.e. quoted above is an underestimate, since it is conditional on the actual panel of sires used by the Institute of Animal Physiology and Genetics Research in both H and L lines. A further conclusion from the magnitude of t_s is that repeated sampling would be useful for increasing the accuracy of a test. The correlation of a test result with breeding value would be increased by 1·23-fold if a test constituted two challenges rather than one and 1·35-fold if three challenges were given.

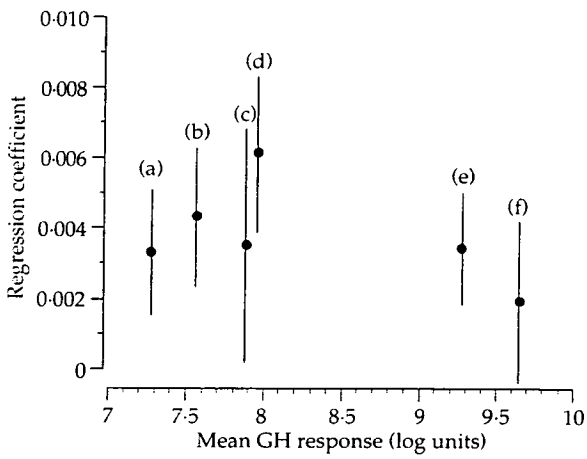


Figure 5 The pooling of regression coefficients for predicted breeding value and mean GH responses from the present experiment (denoted W) and that of Løvendahl *et al.* (1991; denoted L). The results shown are: (a) L, arginine; (b) L, TRH; (c) W, TRH; (d) L, GRF; (e) W, SGRF and DGRF; and (f) L, GRF and TRH administered together.

The final objective was to obtain a greater understanding of the relationship between endogenous GH concentration prior to GRF challenge and response. Given the dramatic results of Tannenbaum and Ling (1984) it is perhaps surprising that no relationship was found between the magnitude of response and the time-interval since the preceding pulse. This time interval is one measure of endogenous GH variation at a given time that might be expected to contain no component between calves. A lack of association was also found between peak concentration of a pulse and the time interval since the previous peak. A further possibility was examined in which the between-calf component of endogenous variation was estimated by the individual grand means over the entire period of serial sampling and the within-calf component was estimated by deviations from a model that included relevant fixed effects and the individual grand means as a covariate. When both these covariates (i.e. within- and between-calf components of error) were included, the importance of within-calf variation in baseline in determining response was clear, yet a regression coefficient of similar magnitude for between-calf variation could not be excluded. If this were the case since more than 0·75 of the variance is found within calves, there are clear benefits obtainable for controlling this component. It remains unclear how this may best be achieved, since the benefit from promoting synchronous pulsing among individuals by some means appears open to question, given the results from regression of response on the time-interval since the preceding pulse.

In summary, the most important findings of this trial have been the confirmation of the positive association of PBV with GH release and the observation of the moderate correlation among repeat challenges either with the same, or different, dose of GRF.

Acknowledgements

The authors would like to thank the staff of Blythbank Farm and Mr H. Bowran and his colleagues at Dryden Farm for their expert care of the calves throughout. The study was supported by the Department of Trade and Industry and the Milk Marketing Board of England and Wales.

References

Agricultural Research Council. 1980 *The nutrient requirement of ruminant livestock.* Commonwealth Agricultural Bureaux, Slough.

Barnes, M. A., Kazmer, G. W., Akers, R. M. and Pearson, R. E. 1985. Influence of selection for milk yield on endogenous hormones and metabolites in Holstein heifers and cows. *Journal of Animal Science* **60:** 271-284.

Bonczek, R. R., Young, C. W., Wheaton, J. E. and Miller, K. P. 1988. Responses of somatotropin, insulin, prolactin and thyroxine to selection for milk yield in Holsteins. *Journal of Dairy Science* **71:** 2470-2479.

Breier, B. H., Bass, J. J., Butler, J. H. and Gluckman, P. D. 1986. The somatotrophic axis in young steers: influence of nutritional status on pulsatile release of growth hormone and circulating concentrations of insulin-like growth factor I. *Journal of Endocrinology* **111:** 209-215.

Diggle, P. J. and Zeger, S. L. 1989. A non-Gaussian model for time series with pulses. *Journal of the American Statistical Association* **84:** 354-359.

Kazmer, G. W., Barnes, M. A., Akers, R. M. and Pearson, R. E. 1986. Effect of genetic selection for milk yield and increased milking frequency on plasma growth hormone and prolactin concentration in Holstein cows. *Journal of Animal Science* **63:** 1220-1227.

Løvendahl, P., Angus, K. D. and Woolliams, J. A. 1991a. The effect of genetic selection for milk yield on the response of growth hormone secretagogues in immature cattle. *Journal of Endocrinology* **128:** 419-424.

Løvendahl, P., Woolliams, J. A. and Sinnett-Smith, P. A. 1992b. Response of growth hormone to various doses of growth hormone releasing factor and thyrotropin releasing hormone administered separately and in combination to dairy calves. *Canadian Journal of Animal Science* **71:** 1045-1052.

Mackenzie, D. D. S., Wilson, G. F., McCutcheon, S. N. and Peterson, S. W. 1988. Plasma metabolite and hormone concentrations as predictors of dairy merit in young Friesian bulls: effect of metabolic challenges and fasting. *Animal Production* **47:** 1-10.

Moseley, W. M., Alaniz, G. R., Clafin, W. H. and Krabill, L. F. 1988. Food intake alters the serum growth hormone response to bovine meal fed Holstein steers. *Journal of Endocrinology* **117:** 253-259.

Patterson, H. D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* **58:** 545-554.

Tannenbaum, G. S. and Ling, N. 1984. The inter relationships of growth hormone (GH)-releasing factor and somatostatin in generation of the ultradian rhythm of GH secretion. *Endocrinology* **115:** 1925-1957.

Tannenbaum, G. S. and Martin, J. B. 1976. Evidence for an endogenous ultradian rhythm governing growth hormone secretion in the rat. *Endocrinology* **98:** 562-570.

Thomas, G. B., Mercer, J. E., Karalis, T., Rao, A., Cummings, J. T. and Clarke, I. J. 1990. Effect of restricted feeding on the concentrations of growth hormone (GH), gonadotropins, and prolactin (PRL) in plasma, and on the amounts of messenger ribonucleic acid for GH, gonadotropin subunits, and PRL in the pituitary of adult ovariectomized ewes. *Endocrinology* **126:** 1361-1367.

Woolliams, J. A. and Smith, C. 1988. The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46:** 333-345.

(*Received* 1 *April* 1992—*Accepted* 3 *August* 1992)

Paper 8

Livestock Production Science, 29 (1991) 1–16
Elsevier Science Publishers B.V., Amsterdam

Physiological attributes of male and juvenile cattle differing in genetic merit for milk yield: a review

J.A. Woolliams[a] and P. Løvendahl[b]
[a]*AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS, UK*
[b]*National Institute of Animal Science, Foulum, Postboks 39, DK-8830 Tjele, Denmark*

(Accepted 17 September 1990)

ABSTRACT

Woolliams, J.A. and Løvendahl, P., 1991. Physiological attributes of male and juvenile cattle differing in genetic merit for milk yield: a review. *Livest. Prod. Sci.*, 29: 1–16.

Experiments to identify physiological attributes of immature cattle that are correlated with breeding value for milk yield have been reviewed. From the results, the measurement of growth hormone and its secretion when challenged with secretagogues currently appear most promising. Although measurements of kinetic parameters associated with thyroid hormones have given positive results, further study has been inhibited by the need to use radioisotopes and by perceived associations with metabolic rate. Measurements of metabolites and insulin under a range of conditions have, in general, not been of significance or, where initial results were encouraging, have proven not to be repeatable. It was concluded that future experimentation must be on a larger scale than has previously been the case and this would best be achieved through international collaboration.

Keywords: dairy cattle; genetics; physiology; juveniles.

INTRODUCTION

The desire to understand the causes of genetic variation in lactation yields of dairy cattle has been fuelled in part by the age- and sex-limited expression of the trait which reduces both the ability to select within the population and the frequency of selection per unit time. The prospect of a physiological understanding of genetic variation offers the possibility that some part of the physiological characteristics, i.e. some of the genes, contributing to genetic superiority can be expressed and measured early in life (preferably pre-pubertal) and in both sexes. These physiological measurements could then be used in breeding schemes. From a theoretical point of view the value of such

indicator traits has been shown by Woolliams and Smith (1988) to be greatest in juveniles and in schemes using multiple ovulation and embryo transfer.

Notwithstanding these practical benefits, this approach offers significant scientific advance. This would be brought about by the unconfounding of the physiological variation and the variation in energy balance, appetite and drive among lactating cows of genetically distinct merit. This avoids the ambiguities of cause and effect inherent in the in vivo study of lactating cows.

Although research in this area has been underway for many years, to the authors' knowledge, findings have rarely been reviewed. This paper is intended to provide a summary of results and problems together with indications of promising directions for future research.

EXPERIMENTAL INFORMATION

A summary of the main experiments reviewed is given in Table 1. A prerequisite for experiments with the objective of understanding the physiology underlying genetic variation is a genetically defined population.

One approach to achieving this is to carry out the physiological studies on bulls that will undergo progeny testing from which precise breeding values will be obtained. The conclusions then depend on the correlation of the study results with progeny test results. The advantages of this method are that results are readily interpreted and are from a population within which the test would be used. The disadvantage of this approach is that for juveniles all experiments necessarily last for 5 years – from the test before puberty to the obtaining of the proof.

A second approach is to study genetically distinct groups of calves achieved through selective breeding. The large divergence in genetic merit that can be obtained between the groups adds power to the experiment even though each individual's breeding value is not accurately known. The advantages of this approach lie in the relatively short time in which an idea can be tested and results obtained. The principal disadvantage is that there is uncertainty as to whether the results hold *within* the groups of high genetic merit for which any test is intended.

Finally, some experiments employ comparisons between breeds such as a dairy breed and a dairy/beef crossbred. The principal reason for using this approach is the belief that differences between the groups will be large and indications can be made from relatively small numbers in a short period without any need to use either of the more structured populations mentioned above. However, interpretation of these trials is difficult: the physiological differences observed may be due to the specialisation of any of the breeds involved – not only the dairy breed, or the difference observed may have been important at some stage of the dairy specialisation but is no longer important for genetic progress *within* the dairy breed. The multibreed approach of Tay-

TABLE 1

Genetic resources of studies reviewed

Reference	Experiment	Breeds[1]	Age	Method of determining merit	Approximate accuracy of genetic determination	Single or divergent population	Numbers[1]
1 Ahlborn-Breier et al. (1987)		NZ Friesian, Jersey	Mature	Progeny test	0.86	Single	48
2 Barnes et al. (1984)		US Holstein	Calves	Pedigree	Unknown	Divergent	7H, 7C
3 Barnes et al. (1985)		US Holstein	4 ages	Pedigree	Unknown	Divergent (10%)	6H, 6C/age
4 Bridges et al. (1987)		NZ Friesian	1 year	Pedigree	Unknown	Divergent (33%)	4H, 4L
5 Graf and Grosser (1979)	1	Fleckvieh	11 months	Progeny test	Unknown	Single	61
	2	Albvieh	11 months	Progeny test	Unknown	Single	15
6 Joakimsen (1975)		Unknown	4–16 months	Progeny test	0.95	Single	94
7 Joakimsen et al. (1971)	1	Finnish Ayrshire	Mature	Progeny test	0.86	Single	20
	2	Norwegian Red	Mature	Progeny test	0.98	Single	22
	3	Swedish Red & White	Mature	Progeny test	0.91–0.98	Single	40
8 Klindt (1988)		US Holstein	Mature	Progeny and Pedigree	0.8	Single	26
9 Land et al. (1983)	1	UK Friesian, HF	6–16 weeks	Pedigree	–	–	6/breed
	2,3	UK Friesian	15 weeks	Pedigree	0.35	Divergent	21H, 21L
10 Løvendahl et al. (1989a)	1,2	Red Danish	3–10 months	Pedigree	0.4	Divergent	38H, 23L
11 Løvendahl et al. (1989b)		UK Friesian	17 weeks	Pedigree	0.4	Divergent	26H, 26L
12 Mackenzie et al. (1988)		NZ Friesian	8 months	Pedigree	0.53–0.62	Divergent (20%)	7H, 6L
13 Osmond (1979)		UK Friesian	12–72 weeks	Pedigree	0.375	Divergent (20%)	48H, 38L
14 Osmond et al. (1981)		UK Friesian	Mature	Progeny test	0.63	Single	152
15 Sejrsen et al. (1984)		Red Danish	15–30 weeks	Pedigree	Unknown	Divergent (20%)	10H, 6L
16 Sinnett-Smith et al. (1987)		UK Friesian	15–24 weeks	Pedigree	0.4	Divergent (15%)	17H, 15L
17 Sorenson et al. (1981)	1	Red Danish	7 months	Progeny test	0.92	Single	25
	2	Danish Black & White	7 months	Progeny test	0.92	Single	53
18 Stark et al. (1978)		UK Friesian	Mature	Progeny test	0.75	Single	119
19 Tilakaratne et al. (1980)		UK Friesian	15 weeks	Pedigree	0.35	Divergent (12%)	21H, 21L
20 Von Flach et al. (1985)		Unknown	Mature	Progeny test	Unknown	Single	318
21 Xing et al. (1988)		NZ Friesian	1–4 weeks	Pedigree	Unknown	Divergent	19H, 20L

[1]H = High, L = Low, C = Control, HF = Hereford × Friesian.

lor and Hnizdo (1987) overcomes some of these problems by reducing the sampling errors from the choice of breeds; however, this has never been directly applied to studies in pre-pubertal calves.

All these approaches run the risk of identifying chance genetic associations present between and within populations arising from genetic drift, and where the physiological difference identified plays no part in the expression of genetic differences in yield. This topic will be discussed later, but it is important to realise that the risks of mis-interpretation are greatest using breed comparisons and least using progeny test results.

RESULTS

Metabolites

Perhaps the simplest approach to identifying predictors is the measurement of energy metabolites such as free fatty acids (FFA), urea (BUN), glucose and ketones such as β-hydroxybutyrate. The justification of this approach has been taken from the observations of Hart et al. (1975, 1978) that genetic differences in propensity to produce milk are associated with genetic variation in the physiology of energy control. This has led to many studies (Tilakaratne et al.; Sjersen et al., 1984; Barnes et al., 1985; Sinnett-Smith et al., 1987; Mackenzie et al., 1988; Xing et al., 1988) examining the levels of these metabolites before and after feeding or during a period of fasting and subsequent refeeding.

Although all metabolites mentioned respond to feeding, there has been no evidence of any difference in this response according to genetic merit. Barnes et al. (1985) noted a tendency for BUN to be greater in calves of high merit than in controls, and Stark et al. (1978) found a positive correlation (0.24 ± 0.11) between BUN of mature bulls and progeny test results. This latter study did not have closely controlled feeding and sampling. However, this finding for BUN has not been a general one in juveniles – other studies do not detect such a difference (e.g. Tilakaratne et al., 1980).

The lack of success obtained from studying metabolites during normal feeding patterns was not repeated when the calves were given the more severe nutritional challenge of fasting. Woolliams and Smith (1988) highlighted the consistent findings of greater concentrations of urea in low merit calves of ages ranging from 4 to 7 months during a period of fasting (Tilakaratne et al., 1980; Sjersen et al., 1984, Sinnett-Smith et al., 1987) although a later publication (Mackenzie et al., 1988) showed no such difference. The results from these studies are analysed in Table 2. The conclusions are certainly more complex than when reviewed by Woolliams and Smith (1988) and, together with further preliminary information from a much larger trial (J.A. Woolliams, in preparation), suggest the co-heritability is much less than the estimate of Woolliams and Smith (1988). Other metabolites studied, in the same

TABLE 2

A summary of results concerning urea concentrations in fasted calves

Reference[1]	Estimated co-heritability[2]	95% confidence interval[3]	Weighting[4]
19	−0.26	(−0.53, −0.07)	76
15	−0.47	(−0.93, −0.22)	32
	−0.17	(−0.47, 0.06)	57
16	−0.26	(−0.48, −0.09)	105
12	0.08	(−0.14, 0.33)	72
Weighted mean	−0.193 ± 0.054		

[1]See Table 1.
[2]Regression coefficient of breeding value on test result (both measured in phenotypic s.d.) assuming a single sample taken at end of fasting period.
[3]Accounts for numbers of calves, numbers of sires and expected genetic divergence in yield (Woolliams and Smith, 1988).
[4]Weighting = [(Confidence interval)/4]$^{-2}$.

TABLE 3

A summary of results concerning FFA concentrations in fasted calves

Reference[1]	Estimated co-heritability	95% confidence interval	Weighting
19	0.18	(0.00, 0.42)	92
15	0	(−0.24, 0.24)	71
	0.17	(−0.14, 0.39)	48
16	−0.07	(−0.20, 0.05)	256
12	−0.12	(−0.34, 0.04)	111
Weighted mean	0.011 ± 0.042		

[1]See Table 2 for explanation of Table headings.

trials, included FFA and β-hydroxybutyrate but these have shown small or inconsistent differences. A summary of findings for FFA is shown in Table 3. This conclusion also holds for glucose except for the study of Xing et al. (1988), the only one to look at glucose in pre-ruminant calves (3 weeks of age), where high breeding index Friesian calves had greater concentrations of glucose than low breeding index calves after they had been fasted overnight.

Thus, although much effort has been put into looking at metabolites during normal feeding and fasting periods, there has been little in the way of consistent results. Perhaps the only possible exception to this, replicated in more than one population, is the differential behaviour of BUN during fasting of calves of differing merit.

Thyroid hormones

Responses in milk yield have been shown to occur when exogenous thyroxine (T_4) is given to dairy cows over short periods (Premachandra and Turner, 1962; Hindery and Turner, 1965), although there are doubts as to whether the responses obtained lead to an increase in full lactation yield (Brumby and Hancock, 1955). These findings and results of Anderson (1971) lead to the hypothesis that thyroxine secretion rates are positively correlated with yield. Further questions then arise as to whether the correlation is in part genetic and whether it is of significant magnitude in the male and at younger ages.

It is the thyroid hormones that have provided the first crucial evidence that genes favourable for milk production may be expressed in the male as a physiological trait, and that appropriate selection in the male using this criterion may increase yield. Tilakaratne et al. (1981) selected male mice at 11 weeks of age (sexually mature) for high (H) and low (L) concentrations of T_4, and found that the 12-day weight of a standardised litter (eight pups) increased in the H compared to the L. The genetic variation in the litter weight of mice at 12 days of age is predominantly attributable to the maternal genotype (Cox et al.,1959; Kidwell and Howard, 1969). It has been inferred (Falconer, 1947) that an important component of mothering ability is milk production, although Hanrahan and Eisen (1970) cast some doubt on this conclusion. Nevertheless, after considering alternative explanations, Tilakaratne et al. (1981) concluded that the selection in the male had influenced the milk yield of the female. Tilakaratne (1979) also presented evidence to suggest that genetic covariances existed between T_4 concentration at 11 weeks of age and younger ages.

This work itself had been stimulated by Joakimsen et al. (1971) who reported significant genetic associations between T_4 degradation rate in mature bulls and their breeding value for fat-correlated milk yield in dairy cows. Their hypothesis was that, in equilibrium, T_4 secretion rates would equal T_4 degradation rates and measurement of the latter was more tractable by observing the disappearance rate of labelled T_4 (Yosef and Johnson, 1967). Their study included data from four groups of bulls of three dairy breeds (see Table 4) and it is notable that the largest and smallest correlations were obtained from the groups with the most accurate proofs (effective number of daughters, $w_d=391$) and least accurate proofs ($w_d=44$) respectively. The authors estimated an upper bound for the heritability of the test (h^2) and pooling the groups estimated a genetic correlation (r_G) of 0.42 ($P<0.01$). Perhaps more appropriately, the pooled value of hr_G was 0.31; this value requires no assumptions about h^2 and is more directly estimated from their data.

Joakimsen (1975), in a separate study, measured T_4 degradation rate and T_3 on 94 younger bulls. Although results were not statistically significant, the association of T_4 degradation rate with breeding value was once again positive in both groups. The importance of T_4 degradation rate was emphasised

TABLE 4

Simple correlations between daily thyroxine degradation in bulls and yearly fat-corrected milk yield of their daughters (from Joakimsen et al., 1971)

AI centre	Breed	No. of bulls	Average no. of daughters/bull	Correlation coefficient
1	Finnish Ayrshire	20	44	0.17
2	Norwegian Red	22	322	0.32
3	Swedish Red & White	20	391	0.47**
4	Swedish Red & White	20	69	0.32
Pooled		82	204	0.30**

*$P<0.05$; **$P<0.01$.

by Sorensen et al. (1981). In this study of T_4 concentration and other kinetic parameters of thyroxine, T_4 degradation rate had the highest genetic correlation with pre-corrected butterfat (0.42). However, this agreement with Joakimsen et al. (1971) is somewhat fortuitous, since as with Joakimsen et al. (1971), the data estimate hr_G more directly than r_G. From the data presented, this value was 0.18 (compared to the 0.31 of Joakimsen et al. (1971)), and is unlikely to have been statistically significant in a study of the size of Sorensen et al. (1981).

One of the disadvantages of measuring T_4 degradation rate is the need to use isotopes, and other authors have tried to simplify matters by correlating milk yield with circulating T_3 or T_4 levels. Osmond et al. (1981) measured T_3 and T_4 in mature Friesian bulls on two occasions, 3 days apart. The concentrations after correction for environmental factors were regressed on progeny test results and the genetic regressions for both T_3 and T_4 concentrations were non-significant (the latter also being negative in sign). Von Flach et al. (1985) found no correlation between either the T_3 or free T_4 concentration of mature bulls and their daughters lactational performance. Graf and Grosser (1979) sampled bull calves at three different ages and found a significant positive association between their progeny test results and T_4 concentrations at 8 months of age but not at later ages. This is then consistent with the results of Osmond et al. (1981) and Von Flach et al. (1985). However, Sorensen et al. (1981) found only a small, negative relationship ($P>0.05$) of T_4 concentrations with progeny test results when sampling was at 7 months of age. Osmond (1979) found no association of T_3 and T_4 with genetic merit in either Friesian or Jersey calves between the ages of 12 and 72 weeks. Seeland et al. (1984) failed to show any significant phenotypic association of T_3 concentration in heifers with subsequent lactation records. In a different approach, Land et al. (1983) compared T_3 and T_4 concentrations of Friesian and Hereford×Friesian bull calves before, during and after feeding and found little or no difference.

In conclusion, there is little evidence that in cattle circulating thyroid hormones in males or young females can act as a predictor of breeding value for dairy merit. There is consistent evidence that the kinetic measurement of T_4 degradation rate does have a positive genetic association with yield, although the magnitude of the genetic regression is uncertain. The reluctance to develop this result stems from (1) the requirement to use radioisotopes and (2) the perceived positive association with maintenance requirements. Standal et al. (1987) investigated this latter possibility and obtained equivocal results. Of relevance to this, Taylor et al. (1986) suggest that an increase in maintenance requirements with yield is inevitable and unavoidable. This remains open to debate.

Growth hormone

Exogenous growth hormone will increase yield (Asimov and Krouge, 1937; Brumby and Hancock, 1955) when administered during lactation both in the short- and long-term. Hart et al. (1978) showed that Friesian cows had greater concentration of GH in their plasma than did Hereford×Friesian cows throughout lactation. More importantly, Kazmer et al. (1986) and Bonczek et al. (1988) found greater circulating GH concentrations in Holstein cows selected for increased yield compared to controls, findings that confirm a trend observed in the much smaller trial of Flux et al. (1984) in New Zealand Friesians. Kazmer et al. (1986) also found that the genetically high-yielding cows were able to secrete more GH in response to a thyrotropin releasing hormone (TRH) challenge than the control cows. Subsequently, Lukes et al. (1989) found that a similar differential release was evident using growth hormone releasing factor (GRF) 45 days after parturition but not at parturition itself. These findings suggest that GH and GH secretion could have a role in defining the physiological attributes of genetic merit for milk yield. However, as outlined in the Introduction, these observations on groups of lactating cows have a concurrent difference in the energy balance due to the differing lactational performance.

When genetically different juvenile cattle were compared, Barnes et al. (1985) found female Holsteins selected for increased yield had greater concentrations of GH than controls at 6 and 12 months of age both before and after feeding, but not at 18 months of age. In accord, Mackenzie et al. (1988) observed greater concentrations in high merit New Zealand Friesian calves both during fasting and after refeeding than in low merit counterparts. Conversely, Land et al. (1983) found that Hereford×Friesian calves had greater GH concentrations than Friesian calves during fasting, and neither Land et al. (1983) nor Woolliams (unpublished results) found any differences before or after normal feeding or during fasting and after refeeding when sampling Friesian calves of differing merit. It is notable that none of the studies looking

at circulating levels in juveniles have tried to account for the pulsatile nature of GH release by using bleeding schedules to estimate baseline, pulse amplitude and pulse frequency. The relative importance of these measures and their association with genetic merit for milk yield is unknown. However, the general conclusion of the studies so far is that circulating GH is unlikely to prove useful or reliable as a predictor.

An alternative approach adopted has been to circumvent the pulsatility and to measure the response of GH to secretagogues. This approach builds on the findings of Kazmer et al. (1986) using TRH in lactating cows, mentioned previously. In three trials, TRH was injected into young calves (Løvendahl et al., 1989a,b; Barnes et al., 1984) with doses ranging from 0.15–0.33 μg/kg and in all trials high-merit calves consistently released more GH than low-merit calves, with statistically significant results in two of the three. Using arginine at a dose rate of 100 mg/kg to release GH, Løvendahl et al. (1989a,b) obtained less convincing results but with the same general trend. However, the release of GH by the pituitary in response to arginine is neither direct nor large and this can complicate interpretation; for example, in the studies of Mackenzie et al. (1988) and Xing et al. (1988) GH concentration decreased following an injection of 40 mg arginine/kg.

In two studies GRF was injected into calves; Massri et al. (1985) gave 0.1 μg/kg to 7-month-old calves and Løvendahl et al. (1989b) gave 0.2 μg/kg to 4-month-old calves; In the latter a clear result was obtained with mean concentrations after injection of 38.6 ng/ml in high-merit compared to 24.0 ng/ml in low-merit calves ($P<0.05$). The results of Massri et al. (1985) are, however, difficult to interpret: the graphs presented indicate a 20% greater response over the 30 min following injection (which includes the peak) in the high-merit compared to low-merit calves, but the analysis of variance favours, if anything, the reverse interpretation. This contradiction arises because the single, maximum GH value obtained for each calf irrespective of time relative to the injection was used in the analysis, a procedure open to considerable sampling variation.

Of the other challenges that modify circulating GH concentration, Land et al. (1983) found no clear discrimination among genetic groups after injection of propionate and, likewise, Osmond (1979) after injection of glucose. Barnes et al. (1984) suggested GH decreased more in high-merit calves after injection of glucose but no data were presented. None of the other challenges, including insulin (Barnes et al., 1985), gave results of significance. Injection of GH itself (Land et al., 1983) showed no difference in the half-life of circulating GH between calves of differing merit.

The pulsatility of GH (and prolactin) in 26 mature bulls was studied by Klindt (1988), and from amongst the various pulsatility parameters a negative relationship was found between GH pulse frequency and dairy merit.

Insulin

Through lactation the plasma insulin concentration decreases initially and rises slowly after peak lactation (Herbein et al., 1985). Low yielding (Hereford×Friesian) crossbred cows had higher insulin concentrations than high yielding cows (Friesians) during lactation (Hart et al., 1978). This finding was confirmed by Bonczek et al. (1988) in Holstein cows where those selected for high yield had the lowest insulin concentrations during peak and mid-lactation. However, during the dry period such differences disappeared (Bines et al., 1983). These findings have not been universal, since Lukes et al. (1989) found similar concentrations among genetic groups throughout lactation, and the smaller studies of Davey et al. (1983) and Barnes et al. (1985) found higher insulin concentrations in cows selected for high yield. Thus there is no clear conclusion to be drawn on the insulin concentrations in lactating cows.

Higher concentrations of plasma insulin were found in crossbred Hereford×Friesian calves than in purebred Friesian calves (Land et al., 1983). Within breed, in calves of selected lines, this finding could not be verified (Land et al., 1983). In the smaller studies of Barnes et al. (1985) and Mackenzie et al. (1988), the highest insulin concentration was associated with high dairy merit. Further, Sinnett-Smith et al. (1987) found no differences in baseline insulin in calves of high and low dairy merit.

In order to study the sensitivity of the pancreas to nutrient stimulation, the insulin response to feeding and intravenous administration of glucose or propionate has been measured. Crossbred (Hereford×Friesian) calves showed a larger insulin response to feeding than did purebred Friesian calves, and similar differences were obtained when insulin release was induced by propionate injection (Land et al., 1983). However, within breed, comparisons of insulin release in high and low dairy merit calves did not reveal significant differences (Land et al., 1983; Sinnett-Smith et al., 1987). Xing et al. (1988) used arginine as an inducer of insulin release in calves but no difference between genetic groups was detected. Although a greater insulin response to glucose was found in high breeding index calves in the study of Mackenzie et al. (1988), an equivalent difference was also observed in the baseline concentrations in that experiment which may have had a contributing influence. The sensitivity of the pancreas to nutrients seems therefore not to be a clear indicator of dairy merit, under the conditions used in these experiments.

The sensitivity of peripheral tissue to insulin is a part of this axis, and may be measured as the change in plasma glucose and free fatty acids following administration of insulin. An increased sensitivity to insulin was found in high breeding index bull calves (Mackenzie et al., 1988) but this was not found in the study of Barnes et al. (1985) nor in the larger study of Land et al. (1983).

On close examination, these findings are largely negative and sometimes

contradictory and may reflect the true circumstances in pre-pubertal animals. It is also possible that this may be a result of uncontrolled interactions of genetic groups with other factors affecting insulin regulation such as feeding regimes (Waghorn et al., 1987) or body condition (McCann and Reimers, 1986). However,induced insulin release has been shown to be a repeatable trait whether induced by propionate (Land et al., 1983) or by folbutamide (Gregory et al., 1982). In the latter study a breed difference (Hereford versus Friesian) was shown at both 12 and 20 months in steers. Thus, although studies have been negative or contradictory, insulin release and sensitivity may deserve further study, but close control of the environmental conditions will be essential.

Other hormones and growth factors

The use of other hormones in attempting to define the attributes of a high merit cow has largely been in the role of challenges. Mackenzie et al. (1988) failed to show any difference in the glucose or free-fatty-acid response to administration of adrenaline. In a small experiment, Bridges et al. (1987) suggested the glucose response in high breeding index heifers was greater after adrenaline injection (1 μg/kg liveweight) and also after 1 μg/kg liveweight of noradrenaline, but not at lower doses. However, Mackenzie et al. (1988) were not able to repeat the result in bull calves, and Sinnett-Smith (unpublished results) found no consistent difference in free fatty acids or glucose after noradrenaline infusions in five-month-old calves, both trials being larger than that of Bridges et al. (1987).

Bolander et al. (1976) compared the circulating placental lactogen concentrations in dairy and beef cows and found 1.7-fold greater concentrations ($P<0.01$) in dairy cows in the final trimester of pregnancy together with some suggestion of a phenotypic correlation of placental lactogen concentration and subsequent lactational performance. This approach is of course limited in its use as a juvenile predictor but it is of interest in the broader context of understanding the physiological attributes of genetically high merit cows.

Allaire et al. (1981) found evidence that steroid concentrations in pre-pubertal heifers were phenotypically associated with their subsequent first-lactation yield but this was not confirmed by Marcus et al. (1987).

Finally there has been one report by Ahlborn-Breier et al. (1987) of a high correlation between progeny test results ($w_d=50$) and IGF1 in the plasma of mature bulls. However, the authors were unable to repeat this result, and the small trial of Mackenzie et al. (1988) was unable to confirm this result in calves.

DISCUSSION

This review has concentrated principally on physiological studies of hormones and metabolites, rather than enzymes. This is because few studies of

genetic divergence in dairy cattle have looked beyond enzyme activities in the blood and the physiological significance of many of these is unclear. Furthermore, in terms of controlling fluxes, the laboratory measurement of a single enzyme in a pathway must be interpreted with caution (Kacser and Burns, 1979). Another, alternative approach of investigating blood or milk polymorphisms has been found previously to give only small or inconsistent associations with milk yield (Neimann-Sorensen and Robertson, 1960; Gonyon et al., 1987; Haenlein et al., 1987) so was not further considered. Associations with milk composition were beyond the scope of the review.

The study of metabolites appears not to be fulfilling its initial promise. Whilst no encouragement was obtained from studies of steady states, the responses of particular metabolites to disturbance away from the steady state gave some hope of success. These were not, in general, confirmed by more intensive study. There remains the possibility that the urea concentration of calves of high and low genetic merit responds differently to fasting. However, it is becoming clear that this association is not as strong as it once appeared and is likely to be too small to yield physiological insights through the study of potential causative factors.

It is the study of hormonal differences that has yielded the most promising results. Of the three hormonal systems that have been studied in most detail — thyroid hormones, insulin and growth hormone — it is the latter for which the evidence is most consistent and the problems appear most tractable. The study of thyroid hormones is hampered by the clear observation that circulating concentrations of T_3 and T_4 are not informative in cattle and that future studies, as in past successes, must involve the kinetics of these hormones. As with insulin, where the picture is confused, the elucidation of an inherent difference in control of the hormones associated with genetic merit for milk production would benefit from a fresh approach.

There does appear to be increasing evidence that the pituitary responsiveness of dairy calves to GH secretagogues may be associated with genetic merit, but this still requires further confirmation. Beyond this, there is a range of questions that would arise leading to a greater physiological understanding of the dairy cow. To answer these questions, more information would be required on the comparative efficacy of the different secretagogues over a range of doses and the relationship of responsiveness to the endogenous pulsatile control of GH in young calves and in dry and lactating cows.

It is clear from the results that this area of research has suffered from a large number of small and uncoordinated studies. Although the problem is physiological, the framework of the study is genetic. It is clear that any association of a physiological process with genetic variation in dairy merit will imply genetic variation in the physiological process. The acknowledgement of the existence of this genetic variation brings with it a need to plan for and control genetic sampling between and within the populations studied. These popula-

tions, by chance alone, can and will often differ simultaneously in dairy merit and some totally unrelated physiological process. Experiments will require many sire groups and offspring to be tested. Even so the statistical significance of a finding in any one trial will still be greater than the significance of the finding in the population as a whole. Thus any physiological association cannot, and should not, be widely accepted until replications are made across population groups. As a result, given the scale of resources required, it is unlikely that practical benefits will accrue without wider co-operation between research groups. This need for replication will bring with it the need for care in controlling environmental conditions under which the testing takes place, since numerous studies have shown that hormonal and metabolic responses are influenced by nutrition before and during testing, age, light intensity and ambient temperature, to name but a few.

REFERENCES

Ahlborn-Breier, G., Breier, B.H., Wickham, B.H. and Gluckman, P.D., 1987. Insulin-like Growth Factor-1 as a measure of genetic merit for milk fat production. Proc. 4th Anim. Sci. Congr. Asian–Aust. Assoc. Anim. Prod. Socs., p. 156.

Allaire, F.R., Bhargava R.K. and Gomes, W.R., 1981. Milk yield and blood serum oestrogens and progesterone in dairy heifers. J. Dairy Sci., 65: 249–258.

Anderson, R.R., 1971. Secretion rates of thyroxin and tri-iodothyronine in dairy cattle. J. Dairy Sci., 54: 1195–1199.

Asimov, G.J. and Krouge, N.K., 1937. The lactogenic preparations from the anterior pituitary and the increase of milk yield in cows. J. Dairy Sci., 20: 289–306.

Barnes, M.A., Kazmer, G.W., Pearson, R.E. and Akers, R.M., 1984. Endocrine and metabolite response after insulin, glucose or TRH in two genetically selected groups of Holstein cattle. J. Dairy Sci., 67: (Suppl.): 161.

Barnes, M.A., Kazmer, G.W., Akers, R.M. and Pearson, R.E., 1985. Influence of selection for milk yield on endogenous hormones and metabolites in Holstein heifers and cows. J. Anim. Sci., 60: 271–284.

Bines, J.A., Hart, I.C. and Morant, S.V., 1983. Endocrine control of energy metabolism in the cow: diurnal variations in the concentrations of hormones and metabolites in the blood plasma of beef and dairy cows. Horm. Metabol. Res., 15: 330–334.

Bolander, F.F., Ulberg, L.C. and Fellos, R.E., 1976. Circulating placental lactogen levels in dairy and beef cattle. Endocrinology, 99: 1273–1278.

Bonczek, R.R., Young, C.W., Wheaton, J.E. and Miller, K.P., 1988. Responses of somatotropin, insulin, prolactin and thyroxine to selection for milk yield in Holsteins. J. Dairy Sci., 71: 2471–2479.

Bridges, J.P., Mackenzie, D.D.S. and Flux, D.S., 1987. Blood metabolite responses to catecholamine injections in heifers of high or low genetic merit for milk fat production. N.Z. J. Agric. Res., 30: 291–296.

Brumby, P.J. and Hancock, J., 1955. The galactopoetic role of growth hormone in dairy cattle. N.Z. J. Sci. Technol. Sect. A, 36: 417–436.

Cox, D.F., Legates, J.E. and Cockerham, C.C., 1959. Maternal influences on body weight. J. Anim. Sci., 18: 519–527.

Davey, A.W.F., Grainger, C., Mackenzie, D.D.S., Flux, D.S., Wilson, G.F., Brookes, I.M. and

Holmes, C.W., 1983. Nutritional and physiological studies of differences between Friesian cows of high and low dairy merit. Proc. N.Z. Soc. Anim. Prod., 43: 67–70.
Falconer, D.S., 1947. Milk production in mice. J. Agric. Sci., 37: 224–225.
Flux, D.S., Mackenzie, D.D.S. and Wilson, G.F., 1984. Plasma metabolite and hormone concentrations in Friesian cows of differing genetic merit measured at two feeding levels. Anim. Prod., 38: 377–384.
Gonyon, D.S., Mather, R.E., Hines, H.C., Haenlein, G.F.W., Arave, C.W. and Grant, S.N., 1987. Associations of bovine blood and milk polymorphisms with lactation traits: Holsteins. J. Dairy Sci., 70: 2585–2598.
Graf, F. and Grosser, I., 1979. Thyroxinwerte von Jungbullen im Bezug zur Eigenleistung und Tochterleistung. Zentralbl. Veterinärmed. Reihe A, 26: 682–686.
Gregory, N.G., Truscott, T.G. and Wood, J.D., 1982. Insulin secretion in relation to fatness in cattle. J. Sci. Food Agric., 33: 276–282.
Haenlein, G.F.W., Gonyon, D.S., Mather, R.E. and Hines, H.C., 1987. Associations of bovine blood groups and milk polymorphisms with lactation traits: Guernseys. J. Dairy Sci., 70: 2599–2609.
Hanranhan, J.P. and Eisen, E.J., 1980. Effect of selection for 12-day litter weight on lactational performance in mice. Aust. J. Biol. Sci., 23: 401–410.
Hart, I.C., Bines, J.A., Balch, C.C. and Cowie, A.T., 1975. Hormone and metabolite differences between lactating beef and dairy cattle. Life Sci., 16: 1285–1292.
Hart, I.C., Bines, J.A., Morant, S.V. and Ridley, J.L., 1978. Endocrine control of energy metabolism in the cow: comparison of the levels of hormones (prolactin, growth hormone, insulin and thyroxine) and metabolites in the plasma of high- and low-yielding cattle at various stages of lactation. J. Endocrinol., 77: 333–345.
Herbein, J.H., Aiello, R.J., Eckler, L.I., Pearson, R.E. and Akers, R.M., 1985. Glucagon, insulin, growth hormone and glucose concentration in blood plasma of lactating dairy cows. J. Dairy Sci., 68: 320–325.
Hindery, G.A. and Turner, C.W., 1965. Effect of administration of L-thyroxine 25 and 50% above secretion rate on lactating cows. J. Dairy Sci., 48: 596–601.
Joakimsen, O., 1975. Estimates of thyroid activity as predictors of breeding value for milk production in cattle. Proc. EAAP 26th Annual Meeting. Warsaw.
Joakimsen, O., Steerberg, K., Lier, H. and Theodorsen, L., 1971. Genetic relationship between thyroid degradation and fat-correlated milk yield in cattle. Acta Agric. Scand., 21: 121–124.
Kacser, H. and Burns, J.A., 1979. Molecular democracy: who shares the controls? Biochem. Soc. Trans., 7:1149–1160.
Kazmer, G.W., Barnes, M.A., Akers, R.M. and Pearson, R.E., 1986. Effect of genetic selection for milk yield and increased milking frequency on plasma growth hormone and prolactin concentration in Holstein cows. J. Anim. Sci., 63: 1220–1227.
Kidwell, J.F. and Howard, A., 1969. The inheritance of growth and form in the mouse. I. Diallel analysis of weight from birth through ten weeks. Growth, 33: 269–289.
Klindt, J., 1988. Relationships among growth hormone and prolactin secretory parameter estimates in Holstein bulls and their predicted differences for lactation traits. J. Anim. Sci., 66: 2784–2790.
Land, R.B., Carr, W.R., Hart, I.C., Osmond, T.J., Thompson, R. and Tilakaratne, N., 1983. Physiological attributes and possible selection criteria for milk production. 3. Plasma hormone concentrations and metabolite and hormonal responses to changes in energy equilibrium. Anim. Prod., 37: 165–178.
Løvendahl, P., Sejrsen, K. and Andersen, B.B., 1989a. Release of growth hormone as a juvenile predictor of dairy merit. In: E. Kalm and T. Liboriussen (Editors), New Selection Schemes in Cattle: Nucleus Programmes. EAAP Publ. No. 44, pp. 134–139.
Løvendahl, P., Woolliams, J.A., Sinnett-Smith, P.A. and Angus, K.D., 1989b. Growth hormone

release in calves selected for high and low dairy merit. In: R.B. Heap, C.G. Prosser and G.E. Lamming (Editors), Biotechnology in Growth Regulation. Butterworths, London.

Lukes, A.J., Barnes, M.A. and Pearson, R.E., 1989. Response to selection for milk yield and metabolic challenges in primiparous dairy cows. Domest. Anim. Endocrinol., 6: 287–298.

Mackenzie, D.D.S., Wilson, G.F., McCutcheon, S.N. and Peterson, S.W., 1988. Plasma metabolite and hormone concentrations as predictors of dairy merit in young Friesian bulls: effect of metabolic challenges and fasting. Anim. Prod., 47: 1–10.

Marcus, G.J., McAllister, A.J. and Hackett, A.J., 1987. Relation of prepubertal serum steroid profiles to first lactation milk yield. Can. J. Anim. Sci., 67: 1021–1026.

Massri, Y.G., Al-Raheem, S.N., Young, C.W. and Wheaton, J.E., 1985. Growth hormone response to growth hormone-releasing factor in selected and control Holstein heifer calves. J. Anim. Sci., 68: 1819–1821.

McCann, J.P. and Reimers, T.J., 1986. Effects of obesity on insulin and glucose metabolism in cyclic heifers. J. Anim. Sci., 62: 772–782.

Neimann-Sorenson, A. and Robertson, A., 1960. The association between blood groups and several production characteristics in three Danish cattle breeds. Acta Agric. Scand., 11: 163–196.

Osmond, T.J., 1979. Studies of hormones in relation to genetic merit of dairy cattle. Ph.D. Thesis, Edinburgh, 99 pp.

Osmond, T.J., Carr, W.R., Hinks, C.J.M., Land, R.B. and Hill, W.G., 1981. Physiological attributes as possible selection criteria for milk production. 2. Plasma insulin, tri-iodothyronine and thyroxine in bulls. Anim. Prod., 32: 159–163.

Premachandra, B.N. and Turner, C.W., 1962. Increased lactation response in cattle with thyroxine 50% above thyroxine secretion rate. J. Dairy Sci., 45:1098–1101.

Seeland, G., Reinecke, P., Schonmuth, G., Segal, L. and Leuthold, G., 1984. Beziehungen zwischen der Schilddrusenaktivität weiblicher Jungrinder und ihrer späteren Milchleistung. 2. Mitteilung: Schätzung genetischer und phänotypischer Parameter der Untersuchungsmerkmale. Arch. Tierzucht, Berlin, 27: 235–245.

Sejrsen, K., Larsen, F. and Andersen, B.B., 1984. Use of plasma hormone and metabolite levels to predict breeding value of young bulls for butterfat production. Anim. Prod., 39: 335–344.

Sinnett-Smith, P.A., Slee, J. and Woolliams, J.A., 1987. Biochemical and physiological responses to metabolic stimuli in Friesian calves of differing genetic merit for milk production. Anim. Prod., 44: 11–19.

Sorensen, M.T., Kruse, V. and Anderson, B.B., 1981. Thyroxine degradation rate in young bulls of Danish dual purpose breeds. Genetic relationship to weight gain, feed conversion and breeding value for butterfat production. Livest. Prod. Sci., 8: 399–406.

Standal, N., Tveit, B. and Svendsen, M., 1987. Thyroid activity in relation to feed intake, growth rate and feed conversion in growing heifers. Domest. Anim. Endocrinol., 4: 201–206.

Stark, A.J., Rowlands, G.J., Manton, R. and McClintock, A.E., 1978. The blood composition of Friesian bulls; variations with age and relationships with results of progeny tests. J. Agric. Sci., Camb, 91: 579–585.

Taylor, St.C.S. and Hnizdo, E., 1987. Multibreed designs. 3. Inter-breed relationships. Anim. Prod., 44: 39–53.

Taylor, St.C.S., Thiessen, R.B. and Murray, J., 1986. Interbreed relationship of maintenance efficiency to milk yield in cattle. Anim. Prod., 43: 37–61.

Tilakaratne, N., 1979. Studies on potential physiological criteria for genetic selection for milk yield. Ph.D. Thesis, University of Edinburgh, 157 pp.

Tilakaratne, N., Alliston, J.C., Carr, W.R., Land, R.B. and Osmond, T.J., 1980. Physiological attributes as possible selection criteria for milk production. 1. Study of metabolites in Friesian calves of high or low genetic merit. Anim. Prod., 30: 327–340.

Tilakaratne, N., Hill, W.G. and Land, R.B., 1981. Direct and correlated responses of mice to selection for the concentration of thyroxine in peripheral plasma. Genet. Res., 38: 157–169.
Von Flach, D., Wassmuth, R. and Dzapo, V., 1985. Stoffwechselparameter als Indikatoren für die Leistungsveranlagung von Rindern. II. Mitteilung: Eignung von Hormonen und Enzymen als Hilfskriterien für die Zuchtwertschatzung. Z. Tierz. Züchtungsbiol., 102: 34–40.
Waghorn, G.C., Flux, D.S. and Ulyatt, M.J., 1987. Effects of dietary protein and energy intakes on growth hormone, insulin, glucose tolerance and fatty acid synthesis in young wether sheep. Anim. Prod., 44: 143–152.
Woolliams, J.A. and Smith, C., 1988. The value of indicator traits in the improvement of dairy cattle. Anim. Prod., 46: 333–345.
Xing, G.Q., Mackenzie, D.D.S., McCutcheon, S.N., Wilson, G.F. and Flux, D.S., 1988. Plasma metabolite and hormone concentrations in Friesian calves differing in genetic potential for milkfat production. N.Z. J. Agric. Res., 31: 159–167.
Yosef, M.K. and Johnson, H.D., 1967. A rapid method for estimation of thyroxine secretion rate of cattle. J. Anim. Sci., 26: 1108–1112.

RESUME

Woolliams, J.A. et Løvendahl, P., 1991. Caracteristiques physiologiques des jeunes bovins et males differant dans leur valeur génétique pour la production laitière: une revue. *Livest. Prod. Sci.*, 29: 1–16 (en anglais).

On passe en revue les expériences visant à trouver chez les jeunes bovins des caractéristiques physiologiques qui soient reliées à leur valeur génétique pour la production laitière. La mesure de l'hormone de croissance et de sa sécrétion paraît être généralement la plus prometteuse. Bien que la mesure des paramètres cinétiques associés aux hormones thyroïdiennes ait donné des résultats positifs, les études supplémentaires ont été freinées à la fois par la nécessité d'utiliser des radio-isotopes et par la conviction que l'utilisation de ces hormones en pratique accroîtrait le rythme métabolique. Le dosage des métabolites et de l'insuline dans diverses situations ne s'est généralement pas montré intéressant ou bien les résultats initiaux encourageants ne se sont pas averés répétables. En conclusion, les expérimentations futures devront être conduites à plus grande échelle que jusqu'ici, ce qui serait le mieux réalisé par une collaboration internationale.

KURZFASSUNG

Woolliams, J.A. und Løvendahl, P., 1991. Physiologische Eigenschaften von männlichen und jungen Rindern mit unterschiedlicher genetischer Veranlagung für die Milchleistung: ein Überblick. *Livest. Prod. Sci.*, 29: 1–16 (auf englisch).

Untersuchungen zur Bestimmung physiologischer Eigenschaften von jungen Rindern, die eine Beziehung zum Zuchtwert für die Milchleistung haben, wurden zusammengefaßt. Von den Ergebnissen erscheint gegenwärtig das Wachstumshormon und seine erzwungene Ausschüttung am Erfolg versprechendsten. Obwohl Messungen von Verlaufsparametern, die mit dem Hormon Thyroxin in Beziehung stehen, positive Ergebnisse erbrachten, sind weitere Studien des Thyroxinhormones wegen der Notwendigkeit, Radioisotopen zu verwenden und wegen der vermuteten Beziehung zum Grundumsatz unterblieben. Ergebnisse von Metaboliten und Insulin unter verschiedenen Bedingungen waren im Allgemeinen nicht signifikant oder konnten, wenn erste Ergebnisse ermutigend waren, nicht wiederholt werden. Die Schlußfolgerung ist, daß weitere Untersuchungen an umfangreicherem Material als bisher notwendig sind, und daß dieses am besten durch eine internationale Zusammenarbeit erreicht werden könnte.

Paper 9

Theor Appl Genet (1990) 80: 503–512

Methods for predicting rates of inbreeding in selected populations

N. R. Wray *, J. A. Woolliams [1] and R. Thompson [1]

University of Edinburgh, Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, UK
[1] Institute of Animal Physiology and Genetics, Edinburgh Research Station, Roslin, Midlothian, EH25 9PS, UK

Received November 15, 1989; Accepted February 15, 1990
Communicated by D. Van Vleck

Summary. In selected populations, families superior for the selected trait are likely to contribute more offspring to the next generation than inferior families and, as a consequence, the rate of inbreeding is likely to be higher in selected populations than in randomly mated populations of the same structure. Methods to predict rates of inbreeding in selected populations are discussed. The method of Burrows based on probabilities of coselection is reappraised in conjunction with the transition matrix method of Woolliams. The method of Latter based on variances and covariances of family size is also examined. These methods are one-generation approaches in the sense that they only account for selective advantage over a single generation, from parents to offspring. Two-generation methods are developed that account for selective advantage over two generations, from grandparent to grandoffspring as well as from parent to offspring. Predictions are compared to results from simulation. The best one-generation method was found to underpredict rates of inbreeding by 10–25%, and the two-generation methods were found to underpredict rates of inbreeding by 9–18%.

Key words: Inbreeding – Effective size – Selection

Introduction

In a random mating population and in the absence of differences in viability and fecundity, all families have equal probabilities of contributing offspring to be parents of the next generation. In a population undergoing selection, families superior for the selected trait will contribute more offspring to the next generation than inferior families and, as a consequence, the rate of inbreeding is higher in selected populations than in randomly mated populations. The parents of superior families are said to confer a selective advantage. The mean level of inbreeding in a given generation t (F_t) and the rate of inbreeding [ΔF, defined as $(F_t - F_{t-1})/(1 - F_{t-1})$] can easily be calculated from pedigree information after selection has occurred, but prediction of inbreeding rate in the planning stage of a breeding programme has proved to be difficult. Frequently, advantages of new breeding schemes are discussed solely in terms of responses to selection, with little regard to the effect of selection on inbreeding, and the assumed rate of inbreeding is appropriate only for random mating populations, therefore making objective assessment of innovations difficult.

Robertson (1961) was the first to discuss the prediction of rates of inbreeding in selected populations of full-sib families (equal numbers of males and females). Although his prediction has been found to severely overpredict rate of inbreeding in simulated populations undergoing intense selection or (mass) selection for traits with high heritabilities (Hill 1985), his approach was pioneering in that it attempted to account for the complete consequences of selective advantage from ancestors to all their descendants. Wray and Thompson (1990) developed the approach of Robertson (1961) to provide formulae of good predicitve value. Their approach is appropriate for populations in which the numbers of males and females differ, although the resulting method is recursive and complex. Therefore, a need exists to examine potentially simpler approaches.

Burrows (1984a) presented a prediction for rate of inbreeding by estimating the probability of pairwise coancestry of selected individuals after a single generation of selection. The method includes higher order terms

* *Present address:* Edinburgh School of Agriculture, West Mains Road, Edinburgh EH9 3JG, UK

that were ignored by Robertson (1961), but it considers only the selective advantage from parent to offspring. Woolliams (1989) presented a method that incorporates the same principles as Burrows (1984a, b) into a transition matrix approach. Using a drift variance argument, Latter (1959) derived an expression for rate of inbreeding in terms of variances and covariances of family size. Although this approach was not derived for the selection case, it has immediate appeal here because of the impact of selection of variances of family size.

We reappraise the method of Burrows (1984a, b) and the method proposed by Woolliams (1989), with particular emphasis on the prediction of coselection of sibs. These methods only consider the proliferation of lineal relatives in a single generation, therefore the development of these to account for proliferation over two generations of selective advantage, i.e. from grandparents to grandoffspring, is considered. A close relationship exists between the transition matrix method and the method of Latter (1959), and a two-generation formula in terms of variances of family size is also derived. Finally, the methods are compared to results from simulations.

Assumptions and notation

Rate of inbreeding is assumed to apply to genes that are neutal with respect to the selected trait and that are unlinked to genes controlling the selected trait. However, if the selected trait is assumed to be controlled by many unlinked loci, each of small additive effect (the infinitesimal model), then the rate of inbreeding at selected loci is expected to be the same as at neutral loci. Selected parents are assumed to be mated at random. Generations are assumed to be nonoverlapping.

M and F denote the total number of males and females available for selection per generation, and m and f (not subscripts) denote the number selected. The letters m and f are used in subscripts in the text as labels for males and females. Let α denote the proportion selected so that $\alpha_m = m/M$ and $\alpha_f = f/F$. Throughout, x and y are used as subscripts for a single sex, where either m or f could be appropriate; similarly, X, Y each represent either M or F, x, y represent m or f and α_x, α_y represent either α_m or α_f. Subscripts i and j are used to describe families and, where both are used together, i and j are assumed to index male and female parents, respectively. Thus m_{ij} denotes the number of males selected from the full-sib family of sire i and dam j. In addition, n_x represents the number of individuals of sex x available from each family for selection. Asymptotic inbreeding rate is denoted by ΔF and effective population size by $N_e = 1/2\,\Delta F$. Finally, $\Phi(a_1)$ denotes $P(w \le a_1)$, where w is any normally distributed random variable and $\Phi(a_1, a_2; \varrho)$ denotes $P(w_1 \le a_1, w_2 \le a_2)$, where w_1 and w_2 are bivariate normal random variables with correlation ϱ.

Prediction of inbreeding by considering the probability of identity by descent

The method of Burrows

Before generalising methods based on identity by descent, it is useful to review the methods of Burrows (1984a, b). The effective population size was defined by $1/N_e = Q_I$, where Q_I is the probability that a pair of genes randomly chosen from distinct, selected individuals (in generation 1 of selection) were contributed by the same inividual of the previous generation. Q_I defined here is equivalent to the Q defined by Burrows (1984b). This definition is equivalent to the inbreeding effective size of Ewens (1982). The inbreeding coefficient of progeny obtained by random mating of the selected individuals is $0.5\,Q_I$ (assuming no existing inbreeding). Burrows (1984a) shows that with full-sib families of fixed size n_x and random mating his definition reduces to that of Wright (1931), which assumes Poisson distribution of family size, $N_e = 4MF/(M+F)$ as $n_x \to \infty$. The condition of large family size is expected, since family size of chosen individuals follows a hypergeometric distribution when available family size (n_x) is constant but, as n_x increases, the Poisson distribution becomes a good approximation to the hypergeometric distribution.

The method of Burrows recognises no sexual dimorphism in the selection process and each selected individual may act as either male or female. This does not adequately account for sex differences in animal selection programmes, and alternative methods to account for this are addressed in the following sections. Q_I depends on the probability of coselection of full and half sibs in the notional gene sampling from among the selected individuals. When indices of relatives are uncorrelated (i.e. when $h^2 = 0$ and family indices are not involved), selection is at random and these probabilities are obtained from the hypergeometric distribution. Let x_{ij} be the number of individuals (of sex x) selected from family (i, j) of size n_x. Then the probability of selecting two full sibs is given by

$$\sum_{i,j} \mathrm{E}\left[\frac{x_{ij}(x_{ij}-1)}{x(x-1)}\right] \tag{1}$$

and

$$\mathrm{E}[x_{ij}(x_{ij}-1)] = \frac{x(x-1)\,n_x(n_x-1)}{X(X-1)}. \tag{2}$$

The probability of selecting two sibs from a common sire is given by

$$\frac{1}{x(x-1)} \sum_i \mathrm{E}\left[\left(\sum_j x_{ij}\right)\left(\sum_j x_{ij}-1\right)\right] \tag{3}$$

$$= \frac{1}{x(x-1)}\left\{\sum_{i,j} \mathrm{E}[x_{ij}(x_{ij}-1) + \sum_i \sum_{j \ne j'} \mathrm{E}[x_{ij}\,x_{ij'}]\right\}.$$

The difference between Eqs. 3 and 1 is the probability of selecting two paternal half sibs. The evaluation thus requires the further result,

$$\mathrm{E}[x_{ij}\,x_{ij'}] = \frac{x(x-1)\,n_x^2}{X(X-1)}. \tag{4}$$

However, when correlations between index values of half sibs (ϱ_{HS}) and of full sibs (ϱ_{FS}) exist then the estimation of Q_I is more complex. If the selection indices of sibs follow a bivariate normal distribution, then Q_I can be approximated by bivariate normal probabilities. The approximation requires the following result, which is more general than that of Burrows (1984a, b), and this generality will be used later when considering two sexes.

Consider a single family containing two groups of size n_1 and n_2. Assign an index G to each individual, G_{kl} ($k = 1, 2$; $l = 1, \ldots n_k$) normally distributed, so that

$$\mathrm{E}[G_{kl}] = 0,\quad \mathrm{E}[G_{kl}^2] = 1,\quad \mathrm{E}[G_{kl}\,G_{kl'}] = \varrho_k$$

and $\mathrm{E}[G_{1l}\,G_{2l'}] = \varrho$, where $0 \le \varrho \le \varrho_k$.

Define T_k as the number of individuals of group k whose index exceeds a threshold a_k. The joint probability generating

function, $M(\theta, \phi)$ of T_1 and T_2 for $p_{uv} = P(T_1 = u;\ T_2 = v)$ is, $M(\theta, \phi) = \sum_1^{n_1} \sum_1^{n_2} p_{uv}\, \theta^u\, \phi^v = E_{z_1, z_2}[\{\Phi(b_1) + \theta[1 - \Phi(b_1)]\}^{n_1}\ \{\Phi(b_2) + \phi[1 - \Phi(b_2)]\}^{n_2}]$, where $b_k = \dfrac{a_k - z_k\, \varrho_k^{1/2}}{(1 - \varrho_k)^{1/2}}$ for $k = 1, 2$ and z_1 and z_2 are distributed as a standardised bivariate normal distribution with correlation coefficient $\varrho_1^{-1/2}\, \varrho\, \varrho_2^{-1/2}$.

The following probabilities can then be derived:

$$
\begin{aligned}
&E[T_k] = n_k(1 - \Phi(a_k)) \\
&E[T_k(T_k - 1)] = n_k(n_k - 1)\,(1 - 2\,\Phi(a_k) + \Phi(a_k, a_k; \varrho_k)) \qquad (5) \\
&E[T_1\ T_2] = n_1\, n_2(1 - \Phi(a_1) - \Phi(a_2) + \Phi(a_1, a_2; \varrho))
\end{aligned}
$$

If the a_k are considered as selection thresholds based on the G_{kl}, it can be seen that the above has some similarities to the selection process. The key difference is that there are no constraints on the sum of the 'T' random variables across the two groups or across a number of families, as there ae in the selection process, where only a predetermined total number are selected. Exact expressions for these quantities can be derived using order statistics, but these are very complex.

To use the formulae 5 in the method of Burrows (1984 a, b), assume the population has a factorial or nested mating structure involving full and half sibs (full-sib family size n_x). Define a common threshold a_x and let the correlation of full-sib family members be ϱ_{FS} and amongst half sibs ϱ_{HS}. ϱ_{HS} can vary according to whether the common parent is the sire of dam, but it is only necessary to consider the former case. Define T_{ij} analogous to the T_k as the number of individuals in family (i, j) greater than the threshold a_x, and apply the following conditions.

(A) for n_x and α_x constant and for all $0 \le \varrho_{HS} \le \varrho_{FS}$, then

$$\frac{1}{x} \sum_{ij} T_{ij} \to \alpha_x \quad \text{as} \quad X \to \infty,$$

(B) as $\varrho_{FS} \to 0$ (and, hence, $\varrho_{HS} \to 0$)

$$E[x_{ij}(x_{ij} - 1)] = k_1\, E[T_{ij}(T_{ij} - 1)]$$

$$E[x_{ij}\, x_{ij'}] = k_2\, E[T_{ij}\, T_{ij'}],$$

which can be used to derive a_x, k_1 and k_2, where k_1 and k_2 are constants.

Substituting Eq. 5 into Eqs. 2 and 4 gives the results of Burrows (1984 b). Formulae 1 and 3 can be used to derive an approximation of coselection probabilities and, hence, Q_I for all ϱ_{HS} and ϱ_{FS}. For example, the probability of coselecting full sibs is

$$\frac{1}{x(x-1)} \sum_{i,j} E[x_{ij}(x_{ij} - 1)] \approx \frac{(n_x - 1)}{\alpha_x^2 (X - 1)} (2\alpha_x + \Phi(a_x, a_x; \varrho_{FS}) - 1) \approx \frac{(2\alpha_x + \Phi(a_x, a_x; \varrho_{FS}) - 1}{s\, \alpha_x^2},$$

for large n_x, where s is the total number of full-sib families.

Methods using transition matrices

Wooliams (1989) presented generalisations of the recurrence relations of Wright (1931) using transition matrices to describe $P_{xy}(t)$, defined as the probability that the genes chosen from distinct selected individuals of sex x and y in generation t are identical by descent. Under three assumptions, the strongest being that the selective advantage of a selected individual over others selected is not inherited in any part by its offspring, then $P_{mf}(t)$, $P_{mm}(t)$ and $P_{ff}(t)$ could be described in terms of $P_{mf}(t-1)$, $P_{mf}(t-2)$, $P_{mm}(t-1)$ and $P_{ff}(t-1)$. An equivalent derivation substitutes $P_{nn}(t-1)$ for $P_{mf}(t-2)$, where $P_{nn}(t-1)$ is the probability that two genes sampled from a single individual of either sex are identical by descent. The coefficients of these relationships were defined by the probabilities that the two individuals sampled had the same maternal or paternal parent (see Woolliams 1989). Then, if $\mathbf{h}_t' = [1 - P_{mm}(t)\ \ 1 - P_{mf}(t)\ \ 1 - P_{ff}(t)\ \ 1 - P_{nn}(t)]$, $\mathbf{h}_t = \mathbf{A}\, \mathbf{h}_{t-1}$ where $\mathbf{A}$ is a matrix of form,

$$\begin{bmatrix} r_1 & 1/2 & r_2 & (1 - 2r_1 - 2r_2)/2 \\ r_3 & 1/2 & r_4 & (1 - 2r_3 - 2r_4)/2 \\ r_5 & 1/2 & r_6 & (1 - 2r_5 - 2r_6)/2 \\ 0 & 1/2 & 0 & 0 \end{bmatrix} \qquad (6)$$

where $0 \le r_i \le 1/4$ for all i. ΔF is described by the behaviour of $P_{mf}(t)$, which is associated with the largest real eigenvalue of $\mathbf{A}(\lambda_{max})$, with $\Delta F \approx (1 - \lambda_{max})$. The equivalence of the matrix $\mathbf{A}$ given in Eq. 6 and that of Woolliams (1989) can be seen by noting that the latter is of the form $\mathbf{D}\,\mathbf{A}\,\mathbf{D}^{-1}$, where $\mathbf{D}$ is a diagonal matrix and thus the two matrices have identical eigenvalues. The existence and bounds of λ_{max} are shown in Appendix I, together with proof that $P_{mf}(t)$ is $O(\lambda_{max})$.

As in the method of Burrows (1984 a, b) the coefficients of transition matrices require probabilities of coselection of full sibs and half sibs when sampling two males, two females and one male and one female. Woolliams (1989) generated these by Monte Carlo simulation, but this gives little indication of their behaviour in general terms. Using Eq. 5 these probabilities can be approximated by bivariate normal probabilities. For (male, male) and (female, female) sampling, the formulae derived in the last section for coselection probabilities of full sibs and half sibs can be used after appropriate substitution of parameters. For (male, female) sampling it is necessary to note that condition (A) defines two thresholds a_m and a_f and to add a further condition, namely,

(C) as ϱ_{FS} (and, hence, $\varrho_{HS} \to 0$ for male and female pairs,

$$E[m_{ij}\, f_{ij}] = k_3\, E[T_{m_{ij}}\, T_{f_{ij}}]$$

$$E[m_{ij}\, f_{ij'}] = k_4\, E[T_{m_{ij}}\, T_{f_{ij'}}].$$

Since the male and female involved in the notional gene sampling do not contribute to the same limit on numbers selected, the moments for $\varrho_{FS} = 0$ are of binomial form rather than hypergeometric, with the result that the constants $k_3 = k_4 = 1$. In addition, a_m and a_f are defined by $\Phi(a_m) = 1 - \alpha_m$ and $\Phi(a_f) = 1 - \alpha_f$. By noting

$$\sum_i E\left[\left(\sum_j m_{ij}\right)\left(\sum_j f_{ij}\right)\right] = \sum_{i,j} E[m_{ij}\, f_{ij}] + \sum_i \sum_{j \ne j'} E[m_{ij}\, f_{ij'}],$$

all coselection probabilities required for constructing the transition matrix can be calculated. For example, the probability of coselection of full sibs when sampling a male and female is given by

$$\frac{1}{m\, f} \sum_{i,j} E[m_{ij}\, f_{ij}] \cong \frac{(\alpha_m + \alpha_f + \Phi(a_m, a_f; \varrho_{FS}) - 1)}{s\, \alpha_m\, \alpha_f},$$

where s is the number of families and ϱ_{FS} is the correlation of indices between male and female full sibs.

The expressions of coselection probabilities derived so far can be simplified. Consider an individual of sex x. The probability of coselecting a relative of given degree of sex y is a product of a function of selection proportions and their index correlation, i.e. $(\alpha_x + \alpha_y + \Phi(a_x, a_y; \varrho) - 1)/(\alpha_x\, \alpha_y)$, and a function of family sizes, where the latter is the ratio of [number of relatives of sex y] and [number of individuals of sex y (excluding self if $x = y$)] where the numbers are those prior to selection. For $\alpha_x = \alpha_y$, the function of the proportions is equal to the reciprocal of the $R(\alpha, \varrho)$ of Burrows (1984 a).

A comparison of Burrows' method and transition matrix methods

The method of the last section explicitly caters for two sexes with differences in their treatment, whilst Burrows' method does not. Woolliams (1989) gave examples which showed that the methods are not, in general, equivalent. Furthermore, in random selection with son replacing sire and daughter replacing dam in a hierarchical mating structure (such as that considered by Gowe et al. 1959), application of Burrows' method gives $\Delta F = 1/8\,M$ when $F \gg M \gg 1$, whilst the method in the previous section gives $\Delta F = 3/32\,M$ as shown by Gowe et al. (1959). Q_I can be written in terms of the elements of matrix **A**, where A_{ij} is the $(i.j)$ element of **A**,

$$\frac{1}{2}Q_I = \frac{1}{2}\left[\frac{M(M-1)}{(M+F)(M+F-1)}A_{14} + \frac{2MF}{(M+F)(M+F-1)}A_{24} + \frac{F(F-1)}{(M+F)(M+F-1)}A_{34}\right].$$

However, when the two sexes are selected on the same index with $n_m = n_f = n_x$, $\alpha_m = \alpha_f = \alpha_x$, and $M = F = X$, the two methods are asymptotically equivalent as $X \to \infty$ for constant α. In these circumstances $P_{mm}(t) = P_{ff}(t)$ and the matrix **A** is of the form,

$$\begin{bmatrix} 1/4-\phi\,\delta/2 & 1/2 & 1/4-\phi\,\delta/2 & \phi\,\delta \\ 1/4-\delta/2 & 1/2 & 1/4-\delta/2 & \delta \\ 1/4-\phi\,\delta/2 & 1/2 & 1/4-\phi\,\delta/2 & \phi\,\delta \\ 0 & 1/2 & 0 & 0 \end{bmatrix}$$

where δ is the probability of sampling genes that were from the same individual of generation $t-1$, when sampling one gene from an individual of each sex in generation t, and $\phi\,\delta$ is the similar probability when sampling one gene from two individuals of the same sex. Since each sex is treated similarly, $\phi \leq 1$, and the characteristic equation of the transition matrix is

$$C(\lambda) = \lambda^4 - (1-\phi\,\delta)\,\lambda^3 - \frac{1}{2}\phi\,\delta\,\lambda^2 + \frac{1}{4}\delta(1-\phi)\,\lambda.$$

Using Newton-Raphson approximation with an initial estimate of 1, for small δ (ignoring terms of δ^2 or higher) a maximum root of $1-(1+\phi)\,\delta/4$ is obtained. This gives $\Delta F = (1+\phi)\,\delta/4$. Using the definition of the section describing Burrows' method $\Delta F = \frac{[(x-1)\,\phi + x]\,\delta}{2(2x-1)}$. Thus, as $x \to \infty$, for constant α, the two definitions become asymptotically equivalent.

Comparison of the Latter-Hill equation and transition matrix methods

Expressions for inbreeding rates in terms of variances and covariances of family size were developed by considering the variance in changes of gene frequency by Latter (1959) for discrete generations and Hill (1972, 1979) for overlapping generations,

$$\Delta F = \frac{1}{32\,ML}\left[2 + \sigma^2_{mm(1)} + 2\left(\frac{M}{F}\right)\sigma_{mm,mf(1)} + \left(\frac{M}{F}\right)^2\sigma^2_{mf(1)}\right] + \frac{1}{32\,FL}\left[2 + \sigma^2_{ff(1)} + 2\left(\frac{F}{M}\right)\sigma_{fm,ff(1)} + \left(\frac{F}{M}\right)^2\sigma^2_{fm(1)}\right], \quad (7)$$

where L is the generation interval, $\sigma^2_{xy(1)}$ is the variance in family size of offspring of sex y contributed to the next generation from parents of sex x and $\sigma_{xm,xf(1)}$ is the covariance in the family size of male and female offspring contributed by parents of sex x. The subscript (1) emphasises that these are variances and covariances taken over a single generation. Equation 7 was derived for the situation where variation in family size is the result of nonheritable causes. In the selection case, the drift variance derivation must be interpreted as the variance in change of gene frequency at a locus neutral with respect to the selected trait.

Variances of family size can be predicted involving arguments similar to those for probabilities of coselection presented earlier. In fact, the elements of the transition matrix defined in Eq. 6 can be written in terms of variances of family size. From drift variance by Eq. 7 in terms of the elements in Eq. 6

$$\Delta F = \frac{1}{8}\left[\left(\frac{M-1}{M}\right)A_{14} + 2A_{24} + \left(\frac{F-1}{F}\right)A_{34}\right].$$

In randomly selected, randomly mated populations of constant size, each generation, the rate of inbreeding calculated from drift variance and identity by descent arguments differ only in second-order terms (Crow and Kimura 1971). However, the formulation given here demonstrates an intrinsic difference in the respective measures. The characteristic equation for the transition matrix treats r_1 (r_6) differently from r_2 (r_5), but in the drift equation above only the sums of $r_1 + r_2$ and $r_5 + r_6$ are important. In nonrandom selection these differences can be of significance. For example, consider a population with M males each mated at random to M females, with each female having one male and M female offspring. If selection rules are imposed such that all M male replacements are chosen from one male parent, and one female replacement is chosen from each female parent, then predicted rates of inbreeding from the transition matrix and Latter-Hill methods are $0.0249 + 0.0520/M$ and $(M+1)^2/32\,M^2$, respectively. If, on the other hand, selection rules are imposed such that one male replacement is chosen from each male, but all female replacements are chosen from a single male, then the two rates of inbreeding are $0.0312 + 0.0665/M$ and $(M+3)/32\,M$. As $M \to \infty$, the Latter-Hill predictions are identical for the two scenarios at 0.0313, but the transition matrix predictions differ by 25% at 0.0249 and 0.0312. This example is perhaps extreme, but it demonstrates that, for selected populations, identity of the transition matrix and Latter-Hill methods are not guaranteed.

Two generation methods

The methods presented so far account for only one consequence of selection on inbreeding, namely, the increased frequency of selecting sibs. The second consequence of selection on inbreeding, namely, the influence of a superior ancestor (older than parent) on the probability that his descendants are selected, is ignored, and these methods are therefore likely to underpredict asymptotic rates of inbreeding. More offspring are likely to be selected from a genetically superior parent than from a genetically average of genetically inferior parent. However, more grandoffspring are likely to be selected from a genetically superior grandparent for two reasons. Firstly, the grandparent was genetically superior as a parent and has already contributed more offspring, and therefore has more grandoffspring available for selection. Secondly, the grandoffspring have inherited superior genes from the grandparent, and so are more likely to be selected than their contemporaries with average or inferior grandparents.

In this section, we attempt to account for the influence of grandparents on the selection of their grandoffspring, as well as the influence of parents on the selection of their offspring. Firstly, the transition matrix method is extended to two generations and, secondly, the variance of family size method of Latter (1959) and Hill (1972, 1979) is extended to include variances of family size over two generations.

Two-generation transition matrix method

Considering the relationships over two generations, individuals can be considered to be full sibs, half sibs, half cousins, full

cousins, double half cousins, one-and-a-half cousins, double full cousins or unrelated at the grandparental level. The probabilities of selecting a gene identical by descent, when sampling two males, two females or a male and a female can be related to the same probabilities in generation t-2 via a transition matrix, **B**, analogous to **A** the one-generation method. ΔF is estimated by $(1-\lambda_{max}^{1/2})$, where λ_{max} is the largest real eigenvalue of **B**. Estimating the elements of **B** is more complex than for **A**, and perhaps is more easily accomplished using simulation. Under random selection $\mathbf{B}=\mathbf{A}^2$.

Two-generation variance of family size method

The drift variance method of Latter (1959) and Hill (1979) can be extended to relate rates of inbreeding to variances of family size from grandparents to grandoffspring, as well as from parents to offspring. The same assumptions as for the one-generation method apply but, in addition, it is assumed that the variance of change in gene frequency over two generations ($V(\delta_{q(2)})$) can be expressed as $V(q_2-q_0)=V(q_2-q_1+q_1-q_0)=V(q_2-q_1)+V(q_1-q_0)\approx 2q(1-q)/2N_e$, where q_t is the gene frequency in generation t and $q_0=q$. Implicit in this assumption is that there is no covariance between changes in gene frequency over the two generations and that $V(q_2-q_1)=V(q_1-q_0)$, which are conditions that may be violated in selected populations.

In the one-generation derivation, variance of change in gene frequency can be attributed to two types of sampling processes, namely, sampling between parents and sampling within heterozygotic parents (the 2's, which are the first elements in the square brackets of Eq. 7, are attributed to the latter cause). In the two-generation derivation, variance of change in gene frequency can be attributed to three types of sampling, sampling between grandparents, sampling within heterozygous grandparents and sampling within heterozygous parents. Details of the derivation are presented in Appendix II and the resulting expression for rate of inbreeding is

$$\begin{aligned}\Delta F = &\frac{1}{256\,ML}\left[\sigma^2_{mm(2)}+2\left(\frac{M}{F}\right)\sigma_{mm,mf(2)}+\left(\frac{M}{F}\right)^2\sigma^2_{mf(2)}\right]\\&+\frac{1}{256\,FL}\left[\sigma^2_{ff(2)}+2\left(\frac{F}{M}\right)\sigma_{fm,ff(2)}+\left(\frac{F}{M}\right)^2\sigma^2_{fm(2)}\right]\\&+\frac{1}{128\,ML}\left[\sigma^2_{mm(1)}+2\left(\frac{M}{F}\right)\sigma_{mm,mf(1)}+\left(\frac{M}{F}\right)^2\sigma^2_{mf(1)}\right]\\&+\frac{1}{128\,FL}\left[\sigma^2_{ff(1)}+2\left(\frac{F}{M}\right)\sigma_{fm,ff(1)}+\left(\frac{F}{M}\right)^2\sigma^2_{fm(1)}\right]\\&+\frac{1}{16\,ML}+\frac{1}{16\,FL},\end{aligned}\qquad(8)$$

where the variance and covariance terms have the same interpretation as in Eq. 7, except that the subscript (2) now represents the variance in family size from grandparents to grandoffspring. For example, $\sigma^2_{mm(2)}$ is the variance in family size of male grandoffspring from male grandparents, which can be written as

$$\sigma^2_{mm(2)}=\sigma^2_{mmm}+\sigma^2_{mfm}+2\sigma_{mmm,mfm},$$

where σ^2_{mmm} and σ^2_{mfm} are the variances in family size from male grandparents to male grandoffspring via male and female offspring, respectively, and $\sigma_{mmm,mfm}$ is the covariance between them. Under random selection and Poisson distribution of family size, Eq. 8, like Eq. 7 reduces to Wright's rate of inbreeding of $1/(8\,ML)+1/(8\,FL)$, and when equal numbers (as far as possible) are chosen from each family, both equations reduce to $3/(32\,ML)+1/(32\,FL)$, as expected (Gowe et al. 1959). The elements of the two-generation transition matrix can also be expressed in terms of variances of family size.

Applications and examples

Parameters

Inbreeding rate will depend on the amount of genetic variation and the covariance of family members, which will affect family size. The methods described reflect this dependence. However, during the selection process such parameters are not constant. The additive-genetic variation will decrease with the initial generations principally through a loss of between-family variation. Provided that inbreeding is not rapid, an equilibrium will be attained in which the loss of between-family variation is regenerated through Mendelian sampling variance originating within families (Bulmer 1971). The genetic variances and covariances are not equal to those in the initial generation nor are they in the same proportions relative to each other. In particular, correlations between relatives decrease and can be as little as half their initial values when h^2 is large. Therefore, there are two potential sets of parameters for use in one-generation formulae, initial and at equilibrium (denoted ϱ_0 and ϱ_e). In the examples, both sets have been used for the one-generation transition matrix method and the method of Burrows (1984a, b), but only the equilibrium parameters have been used for the other methods.

Probabilities of coselection

Simulations were carried out to test the accuracy of the bivariate normal approximations to the coselection probabilities. In the first instance, mass selection was simulated for $M=20$, and $F=20$, 40, 100, 200 with $n_m=n_f=6$. Since the approximation was designed to be exact for index correlations of zero, the simulations used large correlations for demonstration purposes with $\varrho_{HS}=0.6$ and $\varrho_{FS}=0.8$ for both sexes. Simulations are the result of 1,000 realisations of the selection process for each scheme, and the results are shown in Table 1.

The approximation can be seen to have good agreement with the probabilities derived from simulation, with the exception of the probability of selecting two half-sib males when $F=100$ and 200. There are two possible causes for this. Firstly, it may be due to the size of the probability and, secondly, it may be the result of a bias in approximation that occurs when the number of individuals selected is less than the family size, e.g. for $F=100$, 20 males are selected and the half-sib family size for males is 30.

Therefore, a second series of simulations was carried out in which there was only one degree of relationship amongst family members (assumed to be of a single sex) and where the total number of individuals remained constant but family size varied. The selection proportion was kept constant and intraclass correlation was varied. Three thousand realisations were carried out for each scheme and the results are shown in Table 2. It is clear

Table 1. The probabilities of coselection of full and half sibs when sampling selected individuals in a variety of schemes. All assume $M=20$, $n_m=n_f=6$ with $\varrho_{HS}=0.6$ and $\varrho_{FS}=0.8$ for both sexes

		Probability of coselection					
			Full sibs			Half sibs	
		(m,m)[a]	(m,f)[a]	(f,f)[a]	(m,m)	(m,f)	(f,f)
$F=20$	Simulation	0.154	0.178	0.154	0	0	0
	Approximation	0.157	0.186	0.154	0	0	0
$F=40$	Simulation	0.131	0.110	0.077	0.101	0.078	0.067
	Approximation	0.136	0.115	0.078	0.111	0.084	0.070
$F=100$	Simulation	0.102	0.057	0.031	0.237	0.143	0.107
	Approximation	0.115	0.054	0.031	0.328	0.161	0.112
$F=200$	Simulation	0.081	0.027	0.015	0.320	0.170	0.120
	Approximation	0.103	0.028	0.016	0.593	0.199	0.126

[a] (m,m), (m,f) and (f,f) denote notional sampling of two males, one male and one female and two females, respectively

Table 2. The probabilities of coselection of family members when sampling selected individuals in a variety of schemes. Assume a selection proportion of 0.1

Scheme			Correlation among family members			
Number of families (s)	Family size (n_x)		0.25	0.50	0.75	0.90
25	4	Simulation	0.056	0.095	0.152	0.208
		Approximation	0.059	0.098	0.155	0.209
10	10	Simulation	0.161	0.239	0.408	0.576
		Approximation	0.176	0.295	0.466	0.626
4	25	Simulation	0.382	0.525	0.675	0.797
		Approximation	0.469	0.785	1.242	1.669

from Table 2 that the major problem in the approximation is a bias that occurs when selection need only involve one family. The bias is sufficient to give probabilities greater than one. The cause of this bias can be found by considering the derivation of the approximation. Condition (B) considers the limit as $X \to \infty$, whilst n_x and α_x remain constant, and this must imply that the selection process involves many families. This condition becomes untenable when $\varrho \to 1$ and $x \leq n_x$. When $\varrho=1$, the approximation to the probability of coselection for a single sex yields $\frac{(n_x-1)}{(X-1)\,\alpha_x}$ and this exceeds one when $\alpha_x \leq \frac{(n_x-1)}{(X-1)}$.

We put forward a suggestion to overcome this problem, namely, that the value α_x used in the formulae where this problem is encountered is substituted by

$$\alpha'_x=(1-\varrho)\,\alpha_x+\varrho\,\max(\alpha_x,\alpha^*),$$

where α^* is the ratio of the family size, including all relatives of the same or higher degree to that being considered, and the total number of individuals available for selection. A heuristic argument for α^* is that as ϱ approaches one, members of families become less and less distinct, thus selection becomes a two-stage process of, in the first instance, selection of a single family with a selection proportion equal to α^*, followed by a random selection with selection proportion α_x/α^*. As an example consider the scheme in Table 1, with $M=20$, $F=200$, where the probability of coselection of half-sib males is to be estimated. Each individual has 54 male half sibs and 5 male full sibs, thus $\alpha^*=60/1{,}200=0.05$ with $\varrho_{HS}=0.6$, $\alpha_{m'}=0.0367$. Use of α'_m in formulae gives the probability of coselection as 0.345, close to the value of 0.320 estimated by simulation.

Prediction of inbreeding by transition matrix and variance of family size methods

Predictions of rates of inbreeding derived from one- and two-generation transition matrix and variance of family size methods were compared to rates of inbreeding calculated from simulation of populations undergoing mass selection. A range of mating ratios ($M=20$, $F=20$, 40, 100, 200) with three or six offspring of each sex available for selection from each dam was used. An individual's phenotype was simulated as a sum of a breeding value and an environmental effect. The breeding value was sampled from a normal distribution with the mean being the average breeding value of its parents and variance $\frac{1}{2}(1-F)\,\sigma_a^2$, where σ_a^2 is the additive-genetic variance and F is the mean inbreeding coefficient of the parents. The individual environmental component was sampled from a normal distribution with mean zero and variance σ_e^2. A range of heritabilities $\sigma_a^2/(\sigma_a^2+\sigma_e^2)$ was considered: 10^{-6}, 0.1, 0.2, 0.4 or 0.6. Selected parents were mated at random. The rate of inbreeding calculated

Table 3. Rates of inbreeding ($\times 100$) from simulation (ΔF_{sim}) and those predicted using Burrows (1984a, b) (ΔF_B), one and two-generation transition matrix methods ($\Delta F_{1,T}$ and $\Delta F_{2,T}$) and one (Eq. 7) and two (Eq. 8) generation variance of family size equations ($\Delta F_{1,LH}$ and $\Delta F_{2,LH}$)

h^2	ΔF_{sim}	ΔF_B		$\Delta F_{1,T}$		$\Delta F_{2,T}$	$\Delta F_{1,LH}$	$\Delta F_{2,LH}$
		ϱ_0*	ϱ_e**	ϱ_0	ϱ	ϱ_e	ϱ_e	ϱ_e
$M=20$, $F=20$, $n_f=3$								
0.0	1.07	1.05	1.05	1.04	1.04	1.05	1.03	1.03
0.1	1.23	1.11	1.11	1.10	1.10	1.13	1.08	1.11
0.2	1.33	1.18	1.15	1.17	1.14	1.17	1.12	1.16
0.4	1.42	1.31	1.22	1.30	1.21	1.25	1.18	1.24
0.6	1.50	1.45	1.26	1.43	1.25	1.32	1.23	1.31
$M=20$, $F=20$, $n_f=6$								
0.0	1.13	1.15	1.15	1.14	1.14	1.14	1.13	1.13
0.1	1.44	1.29	1.27	1.27	1.25	1.28	1.22	1.27
0.2	1.61	1.42	1.36	1.41	1.34	1.39	1.30	1.38
0.4	1.92	1.72	1.69	1.49	1.45	1.58	1.40	1.57
0.6	1.98	2.04	1.57	2.01	1.55	1.70	1.54	1.68
$M=20$, $F=40$, $n_f=3$								
0.0	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.82
0.1	0.98	0.88	0.87	0.89	0.88	0.89	0.87	0.89
0.2	1.10	0.93	0.91	0.95	0.92	0.95	0.90	0.95
0.4	1.18	1.04	0.96	1.07	0.98	1.03	0.95	1.01
0.6	1.23	1.15	0.99	1.19	1.01	1.08	1.00	1.07
$M=20$, $F=40$, $n_f=6$								
0.0	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.87
0.1	1.17	0.98	0.97	0.98	0.97	1.01	0.96	1.00
0.2	1.30	1.08	1.03	1.09	1.04	1.09	1.02	1.09
0.4	1.50	1.30	1.13	1.33	1.15	1.24	1.12	1.23
0.6	1.50	1.55	1.19	1.59	1.20	1.29	1.17	1.28
$M=20$, $F=100$, $n_f=3$								
0.0	0.71	0.69	0.69	0.71	0.71	0.70	0.70	0.70
0.1	0.82	0.73	0.72	0.76	0.75	0.76	0.74	0.75
0.2	0.94	0.76	0.74	0.81	0.79	0.81	0.77	0.81
0.4	1.05	0.83	0.78	0.93	0.84	0.87	0.82	0.87
0.6	1.02	0.91	0.80	1.05	0.87	0.91	0.85	0.90
$M=20$, $F=100$, $n_f=6$								
0.0	0.73	0.72	0.72	0.72	0.72	0.71	0.70	0.70
0.1	0.95	0.78	0.78	0.81	0.80	0.82	0.78	0.81
0.2	1.10	0.85	0.82	0.90	0.86	0.90	0.83	0.90
0.4	1.24	0.99	0.88	1.10	0.94	1.02	0.92	1.01
0.6	1.23	1.15	0.91	1.35	0.99	1.06	0.96	1.05
$M=20$, $F=200$, $n_f=3$								
0.0	0.66	0.65	0.65	0.66	0.66	0.66	0.66	0.66
0.1	0.82	0.68	0.68	0.72	0.71	0.72	0.70	0.72
0.2	0.86	0.71	0.69	0.77	0.74	0.78	0.73	0.74
0.4	1.00	0.76	0.72	0.89	0.80	0.84	0.79	0.84
0.6	0.99	0.82	0.73	1.02	0.82	0.87	0.81	0.86
$M=20$, $F=200$, $n_f=6$								
0.0	0.87	0.67	0.67	0.67	0.67	0.67	0.67	0.67
0.1	0.87	0.72	0.71	0.75	0.74	0.77	0.73	0.76
0.2	1.03	0.77	0.75	0.84	0.80	0.84	0.78	0.84
0.4	1.12	0.88	0.79	1.03	0.87	0.94	0.85	0.94
0.6	1.19	1.00	0.82	1.24	0.91	0.99	0.89	0.95

* ϱ_0 using initial genetic variances, covariances and correlations prior to selection

** ϱ_e using genetic variances, covariances and correlations at equilibrium

from the simulations was ΔF_{sim},

$$\Delta F_{sim} = \frac{1}{10} \sum_{t=5}^{14} \frac{F_t - \Delta F_{t-1}}{1 - \Delta F_{t-1}},$$

where F_t is the mean level of inbreeding in generation t expected from all possible matings and averaged over 100 replicates. The standard error of F_t increased, with t but did not exceed 0.0030 by generation 14 for any of the examples that are discussed (Table 3). The standard error of $\Delta F_{sim} \times 100$ never exceeded 0.03 for any of the examples. ΔF_{sim} was compared to that expected from theory when $h^2 = 10^{-6}$ (random selection) and to results from other published simulations (Hill 1985; Verrier 1989) and they were found to agree well.

In Table 3 rates of inbreeding calculated in the simulations ΔF_{sim}, are presented along with the predictions calculated from Burrows (1984a, b) ΔF_B; from a one-generation transition matrix approach, $\Delta F_{1,T}$; from the two-generation transition matrix, $\Delta F_{2,T}$; from the Latter-Hill one-generation Eq. 7, $\Delta F_{1,LH}$; and from the two-generation Latter-Hill Eq. 8, $\Delta F_{2,LH}$. The variances of family size required for the Latter-Hill equations were calculated within the simulation described above. The probabilities of coselection necessary for the one- and two-generation transition matrices were calculated as the mean of 1,000 independent simulations.

Under random section ($h^2 = 10^{-6}$), all methods of predicting rates of inbreeding agree well with each other and with ΔF_{sim}. Consider first the predictions using the equilibrium parameters (ϱ_e). There is generally a very good agreement between the one-generation predictions $\Delta F_{1,T}$ and $\Delta F_{1,LH}$ and between the two generation predictions $\Delta F_{2,LH}$ and $\Delta F_{2,T}$ for these populations. The two-generation predictions are always better than the one-generation predictions, but both underestimate the asymptotic rates of inbreeding calculated in the simulation; $\Delta F_{1,T}$ underpredicts ΔF_{sim} by 10–25%, and $\Delta F_{2,T}$ underpredicts ΔF_{sim} by 9–18% in these examples. However, perhaps a more appropriate comparison is the proportion of the observed increase in rate of inbreeding in a selected over a non selected population accounted for the prediction; $\Delta F_{1,T}$ predicts 31–53% $\Delta F_{2,T}$ predicts 44–70% of the total increase in inbreeding. ΔF_B is slightly superior to $\Delta F_{1,T}$ when $M = F$, but becomes considerably inferior as F/M increases, since the different treatment of the sexes becomes important.

Rates of inbreeding calculated using ϱ_0 in some cases appear superior in predicting the true rate of inbreeding to equivalent predictors using ϱ_e. This apparent superiority is, in fact, coincidental; the overestimation of family correlations by ϱ_0 compensates for the incomplete account of the inheritance of selective advantage. Indeed, in some cases, rates of inbreeding calculated using ϱ_0 are greater than the true rate of inbreeding.

Discussion

The significance of the methods presented in this paper may best be discussed by restating the assumptions given by Woolliams (1989) required for $\Delta F_{1,T}$ to equal inbreeding rate (since the methods are closely related, the assumptions will also apply to $\Delta F_{1,LH}$ and $\Delta F_{1,B}$): (i) mating is carried out at random; (ii) genetic variances and covariances remain constant; and (iii) selective advantage relative to others selected is not inherited. Whilst assumption (i) can be satisfied by the breeder, the other two are not, and the results show the importance of both of these.

The importance of assumption (ii) is demonstrated by the discrepancy between the examples using initial and equilibrium parameters. This discrepancy increases as h^2 increases (mass selection), as would be expected from the discrepancy between initial and equilibrium correlations among relatives. However, although initial values are close to true values for high heritabilities, it is more appropriate and natural to use the equilibrium variances since, in doing so, assumption (ii) is then satisfied and there appears to be a more stable relationship with the true inbreeding rate.

Since assumptions (i) and (ii) are satisfied by parameters at equilibrium, the difference between the one-generation models using these and the simulation values is due to breaking assumption (ii), the inheritance of selective advantage since by definition ΔF is a rate independent of previous generations in which genetic variances are equilibrating). The difference is still large: for $h^2 = 0.1$, $\Delta F_{1,T}$ predicts only 30% of the extra inbreeding over random mating, and for $h^2 = 0.6$, only 50% of the extra inbreeding. This trend may be expected from the results of Wray and Thompson (1990), who showed that terms describing inheritance of selective advantage over many generations are functions of $(1 - k\,h^2)$, i.e. of decreasing magnitude as h^2 increases.

The development of the two-stage methods was made in an attempt to introduce the concept of inherited selective advantage into the one-generation methods. Nevertheless, they account for 40–60% of the additional inbreeding through selection, giving only a marginal improvement. It is possible to conceive of n generation methods (ΔF_n) that will converge to ΔF, however, the evidence is that convergence is slow and not obviously predictable from ΔF_1 and ΔF_2. Furthermore, the parameters required for the estimation of ΔF_n will require either extensive simulation or the use of the theory developed by Wray and Thompson (1990). It can be shown (data not presented) that if their formulae for predicting the additional selective advantage from ancestors to descendants are forced to zero for descendants born after one or two generations, then there is good agreement of the resulting estimates of rate of inbreeding with ΔF_1 or ΔF_2, respectively. This has two consequences: firstly, it confirms the cause of the discrepancy between estimates and true rates and, secondly, by examining the results from forcing selective advantage terms to zero after three or more generations, it confirms the slow convergence. As a result, it is unlikely that further development of one- and two-generation methods will be worthwhile.

Verrier et al. (1990) have recently presented a method to predict levels of inbreeding in each generation for populations undergoing selection. The method is an extension of the method of Burrows (1984b), but correctly accounts for sexual dimorphism. The method is recursive, using the level of inbreeding in generation t-1 to predict the rate of inbreeding in generation t and incorporating genetic parameters each generation that account for the effects of selection. However, the method is still 'one-generation' in the sense that it accounts only for the selective advantage of parents to offspring. The asymptotic rate of inbreeding calculated from this method is expected to be equal to the one-generation transition matrix method using ϱ_e.

The accuracy of these one- or two-generation methods may be expected to be better when family indices are used rather than mass selection. This is because, in the latter case, the correlation between relatives has been assumed to be entirely genetic, whereas with family indices it is in part environmental. Indeed, for low h^2 the index correlations of sibs can be principally (but never entirely) of environmental origin. If it were entirely environmental, then the methods $\Delta F_{1,LH}$ and $\Delta F_{1,T}$ are appropriate and unbiased. Thus, there is an a priori case for the bias in estimating additional inbreeding due to selection to be less when using one- and two-generation methods with family indices. Preliminary simulations (N. R. Wray, unpublished results) support this.

An important consideration in the use of the methods of this paper is their flexibility in enabling the modelling of population structure, including overlapping generations. Thus, providing the methods give a reliable 'measure' of ΔF so that the ranking of alternative breeding schemes can be evaluated (as they do in the simulations presented), the methods will have a role in the planning of breeding schemes.

In conclusion, this paper has reviewed and improved various methods for estimating the rate of inbreeding under selection. Whilst these methods are biassed and thus cannot have the reliability of the recursion of Wray and Thompson (1990), they are relatively easy to apply and are likely to give the correct ranking of possible breeding programmes. Therefore, it is valuable to demonstrate the bias with which they predict the asymptotic rate of inbreeding.

Acknowledgements. We would like to thank Prof. W. G. Hill for his critical reading of the manuscript of this paper and for a long series of discussions on this subject.

Appendix I

Eigenvalues and eigenvectors of gene transition matrices

Let $C(\lambda)$ be the characteristic equation of a matrix **A** of the form given in Eq. 6,

$$C(\lambda)=\lambda^4+c_3\lambda^3+c_2\lambda^2+c_1\lambda+c_0,$$

where

$$c_0=\frac{1}{4}(r_1 r_4+r_2 r_5+r_3 r_6-r_1 r_6-r_2 r_3-r_4 r_5)$$

$$c_1=\frac{1}{2}(r_2 r_5-r_1 r_6)+\frac{1}{4}(r_1+r_6-r_3-r_4)$$

$$c_2=\left(r_1 r_6-r_2 r_5+\frac{1}{2}r_1+\frac{1}{2}r_6-\frac{1}{4}\right)$$

$$c_3=-\left(\frac{1}{2}+r_1+r_6\right).$$

Various properties of λ_{max} can be shown. Firstly, $C(1)\geq 0$ with equality if and only if $r_i=1/4$, for $i=1\ldots 6$; furthermore, since $\frac{\partial^j c}{\partial\lambda^j}>0$ for $\lambda=1$, for $j=1\ldots 4$, then $C(\lambda)>0$ for $\lambda>1$; therefore, $\lambda_{max}\leq 1$ should it exist. Secondly, let $k=\frac{1}{4}(1+\sqrt{5})$. $C(k)\leq 0$ with equality if and only if $r_3=r_4=0$ (k is the λ_{max} for full-sib mating). Thus, since $C(1)\geq 0$ and $C(k)\leq 0$, $C(\lambda)$ has at least one root in the interval $k<\lambda\leq 1$; therefore, λ_{max} exists such that $k\leq\lambda_{max}\leq 1$.

Finally, it needs to be shown that $P_{mf}(t)$ is itself of $O(\lambda_{max})$. Three cases can occur:

(i) if $r_3, r_4>0$, then $P_{mf}(t)$ is a linear combination with positive coefficients of the other terms and, since one of these must be $O(\lambda_{max})$, so is $P_{mf}(t)$;

(ii) if $r_3=0$, $r_4\geq 0$, then all individuals are from a common sire and $r_1=r_5=0$, and the result immediately follows as in (i), unless $P_{mm}(t)$ alone is $O(\lambda_{max})$; but this leads to a contradiction, since $P_{mm}(t)$ is a positive linear combination of the others;

(iii) if $r_4=0$, $r_3\geq 0$, then $r_2=r_6=0$ and the result follows from (ii).

This then establishes that $P_{mf}(t)$ is $O(\lambda_{max})$ and $2/(3-\sqrt{5})\leq N_e$. Equality occurs if $r_3=r_4=0$, which is a standard result for full-sib mating. The population size is infinite (it cannot be otherwise) when all selected individuals are unrelated ($r_i=1/4$ for $i=1\ldots 6$) in all generations.

Appendix II

Derivation of the two-generation variance of family size method

The variance of change in gene frequency over two generations, $V(\delta_{q(2)})$ can be expressed as the variance of the mean change in gene frequency from grandparents of either sex to grandoffspring of either sex ($\delta q_{xy(2)}$),

$$V(\delta_{q(2)})=E\left[\frac{1}{4}(\delta q_{mm(2)}+\delta q_{mf(2)}+\delta q_{fm(2)}+\delta q_{ff(2)})\right]^2.$$

Expanding this and omiting terms like $E[\delta q_{mx(2)}][\delta q_{fx(2)}]$ which are zero, it follows that

$$V(\delta_{q(2)})=\frac{1}{16}\{E[\delta^2_{qmm(2)}]+E[\delta^2_{qmf(2)}]+2E[\delta q_{mm(2)}][\delta q_{mf(2)}]\}$$
$$+\frac{1}{16}\{E[\delta^2_{qff(2)}]+E[\delta^2_{qfm(2)}]+2E[\delta q_{ff(2)}][\delta q_{fm(2)}]\}. \quad (A1)$$

The variance of change in gene frequency is attributable to three types of sampling: that due to sampling of genes between grandparents (VGP), that due to sampling of genes within heterozygous grandparents (VHGP) and that due to sampling of genes within heterozygous parents (VHP). Heterozygous parents are generated by the random union of genes from either homozygous or heterozygous grandparents, and so VHP is independent of the other terms. Each term in Eq. A 1 has components due to VGP, VHGP and VHP.

The VGP component of each term in Eq. A 1 is analogous to the second equation on page 501 of Latter (1959) and with $\sigma^2_q=q(1-q)/2$, it follows that

$$E[\delta^2_{qxy(2)}]=\frac{q(1-q)}{2}\frac{\sigma^2_{xy(2)}}{X\mu^2_{xy(2)}},$$

and, similarly,

$$E[\delta q_{xm(2)}][\delta q_{xf(2)}]=\frac{q(1-q)}{2}\frac{\sigma_{xm,xf(2)}}{X\mu_{xm(2)}\mu_{xf(2)}}.$$

It follows that

$$VGP=\frac{q(1-q)}{2}\frac{1}{16M}\left[\frac{\sigma^2_{mm(2)}}{\mu^2_{mm(2)}}+2\frac{\sigma_{mm,mf(2)}}{\mu_{mm(2)}\mu_{mf(2)}}+\frac{\sigma^2_{mf(2)}}{\mu^2_{mf(2)}}\right]$$
$$+\frac{q(1-q)}{2}\frac{1}{16F}\left[\frac{\sigma^2_{ff(2)}}{\mu^2_{ff(2)}}+2\frac{\sigma_{fm,ff(2)}}{\mu_{fm(2)}\mu_{ff(2)}}+\frac{\sigma^2_{fm(2)}}{\mu^2_{fm(2)}}\right].$$

$\mu_{xy(2)}$ can be written as the sum of the mean number of grandoffspring of sex y from grandparents of sex x via male offspring (μ_{xmy}) and the equivalent via female offspring (μ_{xfy}), and is the same whether or not selection is taking place.

$$\mu_{xy(2)}=\mu_{xmy}+\mu_{xfy}=\mu_{xm(1)}\mu_{my(1)}+\mu_{xf(1)}\mu_{fy(1)}$$
$$=\frac{M}{X}\frac{Y}{M}+\frac{F}{X}\frac{Y}{M}=2\frac{Y}{X}.$$

It follows that $\mu^2_{xy(2)}=4\left(\frac{Y}{X}\right)^2$ and

$$VGP=\frac{q(1-q)}{128M}\left[\sigma^2_{mm(2)}+2\left(\frac{M}{F}\right)\sigma_{mm,mf(2)}+\left(\frac{M}{F}\right)^2\sigma^2_{mf(2)}\right]$$
$$+\frac{q(1-q)}{128F}\left[\sigma^2_{ff(2)}+2\left(\frac{F}{M}\right)\sigma_{fm,ff(2)}+\left(\frac{F}{M}\right)^2\sigma^2_{fm(2)}\right].$$

Next, consider the sampling of genes within heterozygotic grandparents (VHGP). The proportion of heterozygotic grandparents is expected to be $2q(1-q)$ and their sampling variance is 1/8 : 1/4 from the binomial variance of the sampling of two genes multiplied by 1/2, since only half the grandoffspring of a given grandparent are expected to receive a gene from that grandparent. The genes from heterozygous grandparents are sampled to form offspring that are homozygous or heterozygous. Each offspring has equal probability of receiving either of the two genes from the grandparent and, since they are random events, there is no sampling covariance between offspring. The sex of the grandparent and the sex of the grandoffspring are therefore irrelevant to the sampling variance. The offspring genes are then sampled as usual to generate the grandoffspring. Each grandoffspring from an offspring family either receives a gene from the grandparent in question or it does not; all those grandoffspring from a given offspring that receive a gene from the grandparent receive the same gene. Therefore, there are $n_{i(xy)}$ grandoffspring of sex y from offspring i of sex x. The sampling covariance summing over all offspring of sex x is $\sum_{i=1}^{X} n^2_{i(xy)}$, which must be averaged overall possible covariances $\left(\sum_{i=1}^{X} n_{i(xy)}\right)^2$. The

VHGP component of Eq. A 1 can then be written as:

$$\mathrm{VHGP}=\frac{q(1-q)}{4}\,\frac{1}{16}\cdot\left[\frac{\sum_{i=1}^{M} n_{i(mm)}^2}{\left(\sum_{i=1}^{M} n_{i(mm)}\right)^2}+2\,\frac{\sum_{i=1}^{M} n_{i(mm)}\,n_{i(mf)}}{\left(\sum_{i=1}^{M} n_{i(mm)}\right)\left(\sum_{i=1}^{M} n_{i(mf)}\right)}+\frac{\sum_{i=1}^{M} n_{i(mf)}^2}{\left(\sum_{i=1}^{M} n_{i(mm)}\right)^2}\right]$$

$$+\frac{q(1-q)}{4}\,\frac{1}{16}\cdot\left[\frac{\sum_{i=1}^{F} n_{i(ff)}^2}{\left(\sum_{i=1}^{F} n_{i(ff)}\right)^2}+2\,\frac{\sum_{i=1}^{F} n_{i(fm)}\,n_{i(ff)}}{\left(\sum_{i=1}^{F} n_{i(fm)}\right)\left(\sum_{i=1}^{F} n_{i(ff)}\right)}+\frac{\sum_{i=1}^{F} n_{i(fm)}^2}{\left(\sum_{i=1}^{F} n_{i(fm)}\right)^2}\right].$$

Notice that the terms in Eq. A 1 relate sex of grandparent to sex of grandoffspring, but terms in Eq. A 2 relate sex of offspring to sex of grandoffspring. Therefore, the terms in the two equations are not directly analogous. Each term in Eq. A 1 has components in several terms in Eq. A 2. The terms in Eq. A2 are found to be directly dependent on one generation means and variances of family size,

$$\frac{\sum_{i=1}^{X} n_{i(xy)}^2}{\left(\sum_{i=1}^{X} n_{i(xy)}\right)^2}=\frac{X\left[\sigma_{xy(1)}^2+\mu_{xy(1)}^2\right]}{X^2\,\mu_{xy(1)}^2}=\left[\frac{1}{X}\left(\frac{X}{Y}\right)^2\sigma_{xy(1)}^2+\frac{1}{X}\right]$$

and, similarly,

$$\frac{\sum_{i=1}^{X} n_{i(xy)}\,n_{i(xf)}}{\left(\sum_{i=1}^{X} n_{i(xm)}\right)\left(\sum_{i=1}^{X} n_{i(xf)}\right)}=\left[\frac{1}{X}\,\frac{X}{F}\,\frac{X}{M}+\frac{1}{X}\right].$$

The $1/X$ terms are due to the sampling of heterozygotic offspring. It follows that

$$\mathrm{VHGP}=\frac{q(1-q)}{4}\,\frac{1}{16\,M}\cdot\left[4+\sigma_{mm(1)}^2+2\left(\frac{M}{F}\right)\sigma_{mm.\,mf(1)}+\left(\frac{M}{F}\right)^2\sigma_{mf(1)}\right]$$

$$+\frac{q(1-q)}{4}\,\frac{1}{16\,F}\cdot\left[4+\sigma_{ff(1)}^2+2\left(\frac{F}{M}\right)\sigma_{fm.\,ff(1)}+\left(\frac{F}{M}\right)^2\sigma_{fm(1)}^2\right].$$

Finally, sampling of heterozygotic parents formed from the random union of genes from grandparents is simply the case of one generation sampling of heterozygotic parents as, before,

$$\mathrm{VHP}=\frac{2q(1-q)}{4}\,\frac{1}{16\,M}\,2+\frac{2q(1-q)}{4}\,\frac{1}{16\,F}\,2$$

$$=\frac{q(1-q)}{16\,M}+\frac{q(1-q)}{16\,F}.$$

It follows that since $\mathrm{V}(\delta_{q(2)})=\mathrm{VGP}+\mathrm{VHGP}+\mathrm{VHP}$ and $\mathrm{V}(\delta_{q(2)})=q(1-q)/N_e$, and $\Delta F=1/(2\,N_e)$ (discrete generations $L=1$), Eq. 8 results. By analogous arguments to Hill (1979), this derivation can be shown to hold for the asymptotic rate of inbreeding in populations with overlapping generations and, hence, each term in Eq. 8 is divided by L.

References

Bulmer MG (1971) The effect of selection on genetic variability. Amer Nat 105:201–211

Burrows PM (1984a) Inbreeding under selection from unrelated families. Biometrics 40:357–366

Burrows PM (1984b) Inbreeding under selection from related families. Biometrics 40:895–906

Crow JF, Kimura M (1971) An introduction to population genetic theory. Harper and Row, New York, Evanston and London

Ewens WJ (1982) On the concept of effective population size. Theor Popul Biol 21:373–378

Gowe RSA, Robertson A, Latter BDH (1959) Environment and poultry breeding problems. 5. The design of poultry control strains. Poult Sci 38:462–471

Hill WG (1972) Effective population size of populations with overlapping generations. Theor Popul Biol 3:278–289

Hill WG (1979) A note on effective population size of populations with overlapping generations. Genetics 92:317–322

Hill WG (1985) Fixation probabilities of mutant genes with artificial selection. Genét Sél Evol 17:351–358

Latter BDH (1959) Genetic sampling in a random mating population of constant size and sex ratio. Aust J Biol Sci 12:500–505

Robertson A (1961) Inbreeding in artificial selection programmes. Genet Res Camb 2:189–194

Verrier E (1989) Prédiction de l'évolution de la variance genétique dans les populations animales d'éffectif limité soumises à selection. PhD thesis, INA-Paris-Grignon, France

Verrier E, Colleau JJ, Foulley JL (1990) Predicting cumulated response to directional selection in finite panmictic populations. Theor Appl Genet 79:833–840

Woolliams JA (1989) Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. Anim Prod 49:1–14

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

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Paper 10

Genet. Res., Camb. (1993), **62**, *pp.* 231–242 *With 3 text-figures*

Prediction of long-term contributions and inbreeding in populations undergoing mass selection

J. A. WOOLLIAMS[1,*], N. R. WRAY[2] AND R. THOMPSON[1]

[1] *AFRC Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, UK*

[2] *Livestock Improvement Unit, Victoria Department of Agriculture, P.O. Box 500, East Melbourne, Victoria 3002, Australia*

(*Received 9 October 1992 and in revised form 26 February 1993*)

Summary

For a population undergoing mass selection, derived from an unselected base population in generation zero, the expected long-term contribution to the population of an ancestor from generation 1 was shown to be equal to that expected during random selection multiplied by $1+i(1-c)^{-1}A_i'$ (where A_i' is one half the breeding value of the ancestor for the trait under selection standardized by the phenotypic standard deviation, i the intensity of selection, and $c=\frac{1}{2}(1-kh_2^2)$ is the competitiveness which is defined by h_2^2 the heritability in generation 2 and k the variance reduction coefficient). It was shown that the rate of inbreeding (ΔF) could be partitioned into three components arising from expected contributions, sampling errors and sampling covariances respectively. Using this result ΔF was derived and shown to be dominated by terms that describe ΔF by variance of family size in a single generation plus a term that accounts for the expected proliferation of lines over generations from superior ancestors in generation 1. The basic prediction of ΔF was given by

$$(1+i^2\rho_m)(8M)^{-1}+(1+i^2\rho_m+2i^2\rho_f)(8F)^{-1}+K[\rho_m(16M)^{-1}+\rho_f(16F)^{-1}]-(8T)^{-1}$$

where M and F are the numbers of breeding males and females, T the number of offspring of each sex, ρ_m and ρ_f are correlations among half-sibs in generation 2 for males and females respectively, and K is a function of the intensity and competitiveness.

1. Introduction

In a novel approach to the prediction of inbreeding, Wray & Thompson (1990) used the concept of the long-term contribution of an ancestor in the first generation of a population and showed that these contributions can be related to the rate of inbreeding. This concept can be described as follows: in a population maintained by the breeding of M males and F females each generation a total of $2^{t-1}(M+F)$ distinct genealogical pathways can be traced back from generation t to generation 1, the long-term contribution of a particular ancestor is $(M+F)$ times the proportion of these pathways that lead back to that ancestor. They developed implicit formulae to relate the long-term contribution of an ancestor to its breeding value when the population was undergoing mass selection, and using these advances they developed a recursive method for the computation of the inbreeding coefficient.

The advantage of such a method was that it took account of the dependence of one generation of selection on previous generations. In random selection the selection processes in each generation proceed independently of all previous generations, but when inheritance is involved then a selective advantage (or disadvantage) of a parent is passed, in part, to its offspring. Consequently the breeding value of the parent has some influence on the selection decisions of all subsequent generations. In mass selection, this influence is mediated entirely through the genes it passes to its offspring.

Only the method of Robertson (1961) had previously allowed for this interdependence of selection decisions. Whilst the method of prediction presented by Wray & Thompson (1990) was considerably better than previous methods (Burrows, 1984*a*, *b*; Verrier, 1989; Wray, Woolliams & Thompson, 1990; Robert-

* Corresponding author.

son, 1961) the method suffered from having no closed form that described the relationship of inbreeding to other predictable genetic parameters.

This paper will derive the explicit relationship between the breeding value of a selected individual in generation 1 and its expected long-term contribution in generation t. It will further derive terms for accumulation of the squared contributions involving both the expected values and chance deviations, together with some adjustments appropriate for small numbers of parents and for when rates of inbreeding are not small. From these an explicit formula for inbreeding is derived, and the origins of its constituent terms identified (i.e. from expected contributions or chance deviations) and their magnitudes evaluated.

Finally it will be shown that the rate of inbreeding is closely approximated by methods based on variances of family size developed by Latter (1959) and Hill (1972) but with a simple correction for the expected inequality of contributions of like-sexed ancestors that arises from the selection process.

2. Notation

Throughout conventions on notation will follow as closely as possible those of Wray & Thompson (1990). Thus the population is propagated through hierarchical random mating of F females with M males ($M \leqslant F$). Each female produces a family of full-sibs of n_f males and n_f females. Each male has $n_m = M^{-1}Fn_f$ offspring of each sex. T is used to denote the total number of offspring of each sex, thus $T = n_m M = n_f F$. X (or Y) or n_x (or n_y) or subscripts x and y are used to specify a single sex either male or female. The long-term genetic contribution from an ancestor i of sex x in generation 1 to descendants of sex y in generation t will be denoted by $r_{i(x)y,t}$ and its expected value by $\mu_{i(x)y,t}$.

The value of $\mu_{i(x)y,t}$ will be shown to be linearly related to $A_{i(x)}$, representing the breeding value of the ancestor for the trait undergoing selection and the slope of the relationship will be given by $\frac{1}{2}b_{xy,t}$. $A_{i(x)}$ has been adjusted so that $E(A_{i(x)}) = 0$. The rate of inbreeding (ΔF) was shown by Wray & Thompson (1990) to be predicted by

$$\tfrac{1}{4}(M+F)^{-2}\left(\sum_{i=1}^{M} r_{i(m)}^2 + \sum_{i=1}^{F} r_{i(f)}^2\right),$$

where $r_{i(x)} = r_{i(x)m,\infty} + r_{i(x)f,\infty}$.

Generation 1 itself was assumed to have been produced by the mating structure described from an unselected base population in which the trait undergoing selection had heritability h_0^2. Generation 1 was the first generation in which selection took place and generation 2 was the first generation produced from selected parents. The heritability in generation t will be denoted h_t^2.

Various parameters relating to the normal distribution will be used throughout: i for the intensity of selection, p for the upper tail probability after truncation at point v with ordinate z, and $k = i(i-v)$ for the variance reduction coefficient. These will be subscripted to refer to particular sexes.

3. Expected long-term contributions

In this section, $\mu_{i(x)y,t}$, the expected long-term genetic contribution from an ancestor i of sex x and with known breeding value, to descendants of sex y after selection in generation t is predicted. Of the $2^{t-1}(M+F)$ pathways, leading back from generation t to generation 1, $2^{t-1}M$ come from males and $2^{t-1}F$ come from females in generation t. Since male and female parents make equal genetic contributions to each individual in each generation exactly half these pathways lead back to males and half to females. Therefore for all descendants of sex y in generation t a total of $2^{t-2}Y$ pathways lead back to each sex in generation 1.

If selection were random each ancestor of sex x would be expected to contribute $2^{t-2}(Y/X)$ pathways from generation 1 to descendants of sex y in generation t. This represents a proportion $2^{t-2}YX^{-1}/2^{t-1}(M+F)$ of all the possible pathways and so the expected long-term contribution of ancestor i of sex x to descendants of sexy y is $(M+F)$ times this proportion i.e. $\mu_{i(x)y,t} = \frac{1}{2}YX^{-1}$. With selection some ancestors are expected to contribute more descendants, and in doing so establish more pathways, than others due to the selective advantage that is a function of the superiority of their breeding value over the breeding value of their contemporaries. It follows that, in a linear model, $\mu_{i(x)y,t} = \frac{1}{2}(YX^{-1} + b_{xy,t}A_{i(x)})$. Denoting the slope by $\frac{1}{2}b_{xy,t}$ allows comparison with Wray & Thompson (1990). $\mu_{i(x)y,t}$ is strictly an expectation that is conditional on $A_{i(x)}$. It follows that the expected number of pathways from i of sex x to descendants of sex y in generation t is $2^{t-1}\mu_{i(x)y,t}$.

The number of pathways from ancestor i of sex x to descendants of sex y in generation t can also be expressed as the sum of pathways to sex y that pass through a male in generation $t-1$ and those that pass through a female in generation $t-1$. Each pathway arriving at an individual of sex w in generation $t-1$ has on average YW^{-1} extensions to sex y in generation t; however, if it is known that the pathway originates from a particular ancestor then the expected number of extensions increases or decreases to some degree according to the breeding value of the ancestor i.e. the expected number of extensions is given by $YW^{-1} + b_{xwy,t}A_{i(x)}$. Therefore, the expected number of pathways to generation t is given by

$$2^{t-1}\mu_{i(x)y,t} = 2^{t-2}\mu_{i(x)m,t-1}(YM^{-1} + b_{xmy,t}A_{i(x)}) + 2^{t-2}\mu_{i(x)f,t-1}(YF^{-1} + b_{xfy,t}A_{i(x)}).$$

If

$$\mu_{i(x)wy,t} = \tfrac{1}{2}(YW^{-1} + b_{xwy,t}A_{i(x)}),$$

then

$$\mu_{i(x)y,t} = \mu_{i(x)m,t-1}\mu_{i(x)my,t} + \mu_{i(x)mf,t-1}\mu_{i(x)fy,t}$$

and this now expresses a recurrence relationship between generations $t-1$ and t.

Note that although linear relationships have been assumed this assumption was tested using simulation by Wray & Thompson (1990) and found to be appropriate.

Wray and Thompson (1990) showed (i) $b_{xy,2} = \frac{1}{2}n_x z_y \sigma_{P2}^{-1}$ where σ_{P2} is the phenotypic standard deviation in generation 2 and (ii) $b_{xwy,t}$ can be approximated by

$$n_w(\tfrac{1}{2})^{t-1}(1-k_w h_{t-1}^2)\left[\prod_{i=2}^{t-2}(1-kh_i^2)\right]z_y\sigma_{P2}^{-1}.$$

In this paper it will be assumed that by generation 2 the values of σ_P^2 and h^2 will be close to their equilibrium values, and so σ_{P2} will be denoted σ_P and $h_2^2 = h_0^2(1-\frac{1}{2}kh_0^2)/(1-\frac{1}{2}kh_0^4)$ will be used in place of h_t^2 for $t \geqslant 2$. It is possible to use the equilibrium values calculated from the equations of Bulmer (1971) but since the coefficients $b_{xwy,t}$ diminish rapidly to zero it is more important to estimate the early coefficients most accurately. If coefficients of competitiveness are defined for each sex by

$$c_x = \tfrac{1}{2}(1-k_x h_2^2) \text{ (and } c = \tfrac{1}{2}(c_m + c_f))$$

then

$$b_{xwy,t} \approx \tfrac{1}{2}n_w c_w c^{t-3} z_y \sigma_P^{-1}.$$

Also by noting $n_x z_y = YX^{-1}i_y$ and defining $A'_{i(x)} = \frac{1}{2}A_{i(x)}\sigma_P^{-1}$ it is observed that

$$\mu_{i(x)y,2} = \tfrac{1}{2}YX^{-1}(1+i_y A'_{i(x)})$$

and

$$\mu_{i(x)wy,t} = \tfrac{1}{2}YW^{-1}(1+i_y c_w c^{t-3}A'_{i(x)}).$$

As t becomes large $\mu_{i(x)wy,t} \to \frac{1}{2}YW^{-1}$ indicating that the influence of ancestor's breeding value on the selection process in generation t decreases to zero.

Let $\boldsymbol{b}_{x,t} = (b_{xm,t}, b_{xf,t})^T$ and $\boldsymbol{z} = (z_m, z_f)^T$, then from collecting terms in the recurrence relationship that are linear in $A_{i(x)}$,

$$\boldsymbol{b}_{x,t} = \boldsymbol{D}\boldsymbol{b}_{x,t-1} + \tfrac{1}{2}X^{-1}[M(b_{xmm,t}, b_{xmf,t})^T + F(b_{xfm,t}, b_{xff,t})^T],$$

where $\boldsymbol{D}$ is the matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2}MF^{-1} \\ \frac{1}{2}FM^{-1} & \frac{1}{2} \end{pmatrix}.$$

$\boldsymbol{D}$ describes the expected dispersion of genes through the population in the absence of selection from generation to generation. For example, $\frac{1}{2}MF^{-1}$ is the expected number of copies of a gene sampled from selected females in generation t^{-1}, among the selected males of generation t. $\boldsymbol{D}$ is idempotent, i.e. $\boldsymbol{D}^2 = \boldsymbol{D}$, and this property embodies the phenomenon that only a single generation is required to disperse genes from one sex through a homogeneous, random mating, diploid population with discrete generations.

Using the property of idempotence Appendix 1 shows that the recurrence relationship can be solved to give

$$\boldsymbol{b}_{x,t} = \tfrac{1}{2}\sigma_P^{-1}[S_{t-3}\, i(MX^{-1}, FX^{-1})^T + c^{t-2}(i_m MX^{-1}, i_f FX^{-1})^T],$$

and consequently

$$\mu_{i(x)y,t} = \tfrac{1}{2}YX^{-1}[1 + A'_{i(x)}(S_{t-3}\, i + c^{t-2}i_y)]$$

and

$$\mu_{i(x)y,\infty} = \tfrac{1}{2}YX^{-1}[1 + iS_\infty A'_{i(x)}]$$

where

$$S_t = \sum_{i=0}^{t} c^i,$$

and

$$S_\infty = \sum_{i=0}^{\infty} c^i = (1-c)^{-1} = 2(1+kh_2^2)^{-1}.$$

Thus $b_{xy,\infty}/b_{xy,2} = i/i_y(1+kh_2^2)$ independent of x. Application of this formula to the results of Wray & Thompson (1990) shows accurate prediction: some comparisons can be made with their simulation results present in their Table 4 which differ by $\leqslant 0{\cdot}01$ from the expectation given above.

4. Prediction of the rate of inbreeding

Wray & Thompson (1990) showed that ΔF is related to the expectation of the squared contributions and this involves not only the expectation of the square of the conditional expectations calculated in the previous section but also the expectation of the conditional variance. In this section the methods required to derive these expectations are described. From the definition,

$$\Delta F = \tfrac{1}{4}(M+F)^{-2}\left(\sum_{i=1}^{M} r_{i(m)}^2 + \sum_{i=1}^{F} r_{i(f)}^2\right),$$

thus

$$E(\Delta F) = \tfrac{1}{4}(M+F)^{-2}\left(\sum_{i=1}^{M} E(r_{i(m)}^2) + \sum_{i=1}^{F} E(r_{i(f)}^2)\right)$$
$$= \tfrac{1}{4}(M+F)^{-2}(ME(r_{i(m)}^2) + FE(r_{i(f)}^2)).$$

Suppose the squared contribution can be decomposed into s elements so that

$$r_{i(x)}^2 = \sum_{j=1}^{s} r_{i(x)}^2(j),$$

then

$$\Delta F = \tfrac{1}{4}(M+F)^{-2}\sum_{j=1}^{s}\sum_{x=m,f} XE(r_{i(x)}^2(j)).$$

The expression has been given in this form as it will be seen that simplification of terms is derived from it.

Wray & Thompson (1990) derive ΔF in a 'lateral' and recursive accumulation: terms involved in $r^2_{i(x),2}$ are derived and accumulated; the expected change in contribution in these moving from generations 2 to 3 is derived and $r^2_{i(x),3}$ is derived by adding further terms originating in generation 3, the recursion then proceeds through generations until contributions become negligibly small.

Following equations (24) and (27) of Wray & Thompson (1990)

$$E(r^2_{i(x),2}) = E\{\mathbf{1}^T[\boldsymbol{\mu}_{i(x),2}\boldsymbol{\mu}^T_{i(x),2} + \boldsymbol{V}_{i(x),2} + \boldsymbol{C}_{i(x),2}]\mathbf{1}\}$$

where $\mathbf{1}^T = (1,1)$ and $\boldsymbol{V}_{i(x),t}$ and $\boldsymbol{C}_{i(x),t}$ are the matrices of new contributions arising in generation t from binomial sampling and from additional covariance through co-selection of sibs (i.e. the squared contribution is the sum of the mean contribution squared plus the variance of the contribution).

In generation 3

$$E(r^2_{i(x),3}) = \mathbf{1}^T[\boldsymbol{M}_{i(x),3}(\boldsymbol{\mu}_{i(x),2}\boldsymbol{\mu}^T_{i(x),2} + \boldsymbol{V}_{i(x),2} + \boldsymbol{C}_{i(x),2})\boldsymbol{M}^T_{i(x),3} + \boldsymbol{V}_{i(x),3} + \boldsymbol{C}_{i(x),3}]\mathbf{1},$$

where

$$\boldsymbol{M}_{i(x),t} = \begin{pmatrix} \mu_{i(x)mm,t} & \mu_{i(x)fm,t} \\ \mu_{i(x)mf,t} & \mu_{i(x)ff,t} \end{pmatrix}$$

and ultimately

$$E(r^2_{i(x)}) = \mathbf{1}^T\boldsymbol{P}_{i(x),3}\boldsymbol{\mu}_{i(x),2}\boldsymbol{\mu}^T_{i(x),2}\boldsymbol{P}^T_{i(x),3}\mathbf{1} + \sum_{t=2}^{\infty}\mathbf{1}^T\boldsymbol{P}_{i(x),t+1}\boldsymbol{V}_{i(x),t}\boldsymbol{P}^T_{i(x),t+1}\mathbf{1} + \sum_{t=2}^{\infty}\mathbf{1}^T\boldsymbol{P}_{i(x),t+1}\boldsymbol{C}_{i(x),t}\boldsymbol{P}^T_{i(x),t+1}\mathbf{1},$$

where

$$\boldsymbol{P}_{i(x),t} = \prod_{j=t}^{\infty}\boldsymbol{M}_{i(x),j}.$$

Thus $E(r^2_{i(x)})$ has three terms: due to the squared mean, to binomial sampling and to co-selection of sibs. By deriving these terms separately and accumulating them vertically over generations a closed expression for ΔF is obtained.

To achieve the vertical accumulation it is necessary to derive $\mathbb{P}_{i(x),t+1}$ explicitly. Appendix 2 shows $\mathbb{P}_{i(x),t+1}$ to be

$$\mathbb{D} + A'_{i(x)}c^{t-2}i\mathbb{D}\mathbb{\Theta} + A'^2_{i(x)}(c^{t-2})^2.icq\,\mathbb{D}\mathbb{\Lambda} + \text{terms in } A'^3_{i(x)}.$$

Here $\mathbb{\Theta}$ and $\mathbb{\Lambda}$ are diagonal matrices whose elements, along with q, depend on the intensity and competitiveness of the selection in each sex.

A general form for each of $\boldsymbol{V}_{i(x),t}$ and $\boldsymbol{C}_{i(x),t}$ is $u_{0,t}\boldsymbol{\delta} + u_{1,t}A'_{i(x)}\boldsymbol{\varepsilon} + u_{2,t}A'^2_{i(x)}\boldsymbol{\zeta}$ where $\boldsymbol{\delta}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\zeta}$ are matrices that depend on genetic parameters but are independent of t. $\boldsymbol{P}_{i(x),t+1}$ is of the form $\boldsymbol{\alpha} + c^{t-2}A'_{i(x)}\boldsymbol{\beta} + c^{2t-4}A'^2_{i(x)}\boldsymbol{\gamma}$.

Since $E(A'_{i(x)}) = 0$ and contributions of $O(A'^3_{i(x)})$ or higher are ignored the contributions to $\frac{1}{4}(M+F)^{-2}XE(r^2_{i(x)})$ are of the form $v_0 + E(A'^2_{i(x)})\Sigma^4_{j=1}v_j$, where

$$v_0 = (\Sigma u_{0,t})(\mathbf{1}^T\boldsymbol{\alpha}\boldsymbol{\delta}\boldsymbol{\alpha}^T\mathbf{1}), \quad v_1 = (\Sigma u_{2,t})(\mathbf{1}^T\boldsymbol{\alpha}\boldsymbol{\zeta}\boldsymbol{\alpha}^T\mathbf{1}),$$
$$v_2 = (\Sigma c^{2t-4}u_{0,t})(\mathbf{1}^T\boldsymbol{\beta}\boldsymbol{\delta}\boldsymbol{\beta}^T\mathbf{1}),$$
$$v_3 = (2\Sigma c^{t-2}u_{1,t})(\mathbf{1}^T\boldsymbol{\beta}\boldsymbol{\varepsilon}\boldsymbol{\alpha}^T\mathbf{1}),$$
$$v_4 = (2\Sigma c^{2t-4}u_{0,t})(\mathbf{1}^T\boldsymbol{\gamma}\boldsymbol{\delta}\boldsymbol{\alpha}^T\mathbf{1}).$$

The infinite sums in these expressions are convergent since $u_{0,t}$, $u_{1,t}$ and $u_{2,t}$ are $O(2^{-t})$. The symbols v_0 and v_j will be referred to in Appendix 3 in order to aid the identification of the origin of terms. $E(A'^2_{i(x)})$ is the correlation between half-sibs with common parent of sex x in generation 2 prior to selection and will be denoted ρ_x and his value $\frac{1}{4}h^2_0(1-k_x h^2_0)/(1-\frac{1}{2}kh^4_0)$.

The matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in $\boldsymbol{P}_{i(x),t+1}$ have already been derived and are all of the form $\boldsymbol{D}$ post-multiplied by some diagonal matrix. Therefore the general form for contributions to $E(r^2_{i(x)})$ is always $\frac{1}{4}(M+F)^{-2}\mathbf{1}^T\boldsymbol{D}\boldsymbol{T}\boldsymbol{D}^T\mathbf{1}$ for some matrix $\boldsymbol{T}$. If $\boldsymbol{T}$ has elements t_{ij}, then

$$\tfrac{1}{4}(M+F)^{-2}\mathbf{1}^T\boldsymbol{D}\boldsymbol{T}\boldsymbol{D}^T\mathbf{1} = (F^2t_{11} + MFt_{12} + MFt_{21} + M^2t_{22})/(16M^2F^2).$$

(i) *Squared mean contributions*

For ancestors of sex x the contribution to ΔF is $\frac{1}{4}X(M+F)^{-2}\mathbf{1}^T\boldsymbol{P}_{i(x),3}\boldsymbol{\mu}_{i(x),2}\boldsymbol{\mu}^T_{i(x),2}\boldsymbol{P}^T_{i(x),3}\mathbf{1}$. Since

$$\mu_{i(x)w,2}\mu_{i(x)y,2} = \tfrac{1}{4}WYX^{-2}(1 + A'_{i(x)}(i_w + i_y) + A'^2_{i(x)}i_w i_y),$$
$$X\boldsymbol{\mu}_{i(x),2}\boldsymbol{\mu}^T_{i(x),2} = \tfrac{1}{4}X^{-1}\left[\begin{pmatrix} M^2 & MF \\ MF & F^2 \end{pmatrix} + A'_{i(x)}\begin{pmatrix} 2M^2i_m & MF(i_m+i_f) \\ MF(i_m+i_f) & 2F^2i_f \end{pmatrix} + A'^2_{i(x)}\begin{pmatrix} M^2i^2_m & MFi_m i_f \\ MFi_m i_f & Fi^2_f \end{pmatrix}\right].$$

Contributions are given in Appendix 3.

(ii) *Binomial sampling*

In generation 2, if success is taken as a selected offspring of sex y the binomial parameter for ancestor i of sex x is given by equation (22) of Wray & Thompson (1990) as

$$p_{i(x)y,2} = p_y + n_x^{-1}b_{xy,2}A_{i(x)} = p_y(1 + i_y A'_{i(x)}).$$

The variance assuming independent trials is therefore $n_x p_{i(x)y,2}(1-p_{i(x)y,2})$. The contribution to $E(r^2_{i(x),2})$ is

$$\boldsymbol{V}_{i(x),2} = \tfrac{1}{4}X^{-1}T[\mathbb{V}_0 + A'_{i(x)}\mathbb{V}_1 + A'^2_{i(x)}\mathbb{V}_2]$$

where $\mathbb{V}_0$, $\mathbb{V}_1$ and $\mathbb{V}_2$ are diagonal matrices:

$$\mathbb{V}_0 = \text{diag}(p_m(1-p_m), p_f(1-p_f)),$$
$$\mathbb{V}_1 = \text{diag}((1-2p_m)z_m, (1-2p_f)z_f)$$

and

$$\mathbb{V}_2 = \text{diag}(-z^2_m, -z^2_f).$$

In subsequent generations, new contributions arise from binomial sampling of gene pathways leading to offspring of sex y in generation t from parents of sex w. The probability of success is given by $p_{i(x)wy,t} = p_y(1+i_y c_w c^{t-3} A'_{i(x)})$ and each parent has a family size of n_w. The expected contribution of these variances is weighted by the expected number of ancestors $\mu_{i(x)w,t-1}$. The contributions are then summed over w. Thus

$$V_{i(x),t} = 2^{-t}X^{-1}T[V_0 + A'_{i(x)}[iS_{t-3}V_0 + c^{t-2}V_1] + A'^2_{i(x)}[(ic^{t-3}S_{t-4} + gc^{2t-6})V_1 + \tfrac{1}{2}(c_m^2 + c_f^2)c^{2t-6}V_2]]$$

$V_{i(x),t}$ can be thus separated into 5 components denoting terms in V_0, $A'_{i(x)}V_0$, $A'_{i(x)}V_1$, $A'^2_{i(x)}V_1$ and $A'^2_{i(x)}V_2$. The resulting terms are listed in Appendix 3.

(iii) *Co-selection of sibs*

The original description of co-selection by Wray & Thompson (1990) was incomplete, leading to the omission of important contributions from the co-selection of sibs, including half-sibs. The omissions can contribute up to 8 % of the total for the examples of mass selection considered here and can be many times larger than the terms considered by Wray & Thompson (1990). The original description was also incorrect in using $h^2/4$ as a correlation between half-sibs which can lead to important errors as h^2 increases.

The objective is to estimate the variance of family size conditional on the breeding value of an ancestor in generation 1. Co-selection occurs through the covariances of selection probabilities of full- or half-sibs that are not accounted for by regression on the breeding value of the ancestor in generation 1. It is useful to split the consideration of this into two parts; generation 2 and generation 3 onwards, since in generation 2 the ancestor is in fact the parent.

In generation 2, the covariance between half-sibs or full-sibs, arising from the parent of sex x are already fully accounted for in the term describing expected contributions, leaving only those arising in full-sib families through the parent of sex x' (i.e. the sex other than x). The probability of selecting an individual of sex y with parents of breeding value $A_{i(x)}$ and $A_{j(x')}$ is given by $p_y(1+i_y A'_{i(x)} + i_y A'_{j(x')})$. Thus when mating is at random, and conditional on the breeding value $A_{i(x)}$, the covariance of the selection probabilities of two full-sibs of sex y and z is given by

$$E[p_y p_z(1+i_y A'_{i(x)} + i_y A'_{j(x')})(1+i_z A'_{i(x)} + i_z A'_{j(x')})] - p_y p_z(1-i_y A'_{i(x)})(1+i_z A'_{i(x)}) = p_y p_z i_y i_z \rho_{x'}.$$

This covariance will arise between all full-sib pairs for ancestor $i(x)$, i.e. $X^{-1}T(n_f-1)$ times if $y=z$ and $X^{-1}Tn_f$ if $y \neq z$. For simplicity of exposition the (n_f-1) will be treated as n_f at present. Thus $\boldsymbol{C}_{i(x),2} = \frac{1}{4}X^{-1}Tn_f\rho_{x'}\boldsymbol{C}_0$ where $\boldsymbol{C}_0 = (p_m i_m, p_f i_f)^T(p_m i_m, p_f i_f)$. The terms arising from $\boldsymbol{C}_{i(x),2}$ are given in Appendix 3 for a single sex x.

In generation t where $t \geqslant 3$, for an offspring of a parent of sex w of breeding value A_w with ancestor $i(x)$ in generation 1 and mated at random to an individual of sex w' with breeding value $A_{w'}$ the probability of selection expressed as a regression on $A_{i(x)}$, A_w and $A_{w'}$ is

$$p_y(1 + c^{t-3}c_w i_y A'_{i(x)} + i_y(A'_w - c^{t-3}c_w A'_{i(x)}) + i_y A'_{w'}).$$

Thus conditional on the ancestor's breeding value, with random mating the covariance not only arises from w' but from w as well. For two full-sibs of sexes y and z the covariance is

$$p_y p_z i_y i_z[(\rho_{tw} + \rho_{w'}) - c_w^2 c^{2t-6}\rho_x];$$

but for two half-sibs of sexes y and z with common male parent there is also a covariance of the form $p_y p_z i_y i_z[\rho_m - c_m^2 c^{2t-6}\rho_x]$. Strictly speaking in generation $t \geqslant 3$ the correlation of half-sibs with common parent of sex w is approximately $\frac{1}{2}h_2^2 c_w$, but it will be assumed that for the cases considered this differs little from ρ_{tw}, the correlation of half-sibs in generation 2. With the hierarchical mating structure considered here only male half-sibs are formed.

There are $M^{-1}Tn_f(d-1)$ half-sib pairs in a male parent's family where $d = FM^{-1}$. Thus for a male descendant in generation $t-1$ where $t \geqslant 3$

$$\boldsymbol{C}_{i(x)m,t} = 2^{-t}M^{-1}T(n_f(\rho_m + \rho_f) + n_f(d-1)\rho_m - n_m c_m^2 c^{2t-6}\rho_x)\boldsymbol{C}_0,$$

whereas for a female descendant

$$\boldsymbol{C}_{i(x)f,t} = 2^{-t}F^{-1}T(n_f(\rho_m + \rho_f) - n_f c_f^2 c^{2t-6}\rho_x)\boldsymbol{C}_0.$$

Following the procedure set out in the section on binomial sampling the terms require weighting by $\mu_{i(x)w,t-1}$ and summed for $w = m$ and f. The terms arising are listed in Appendix 3. Terms in $(c^3)^t$ have been ignored and also simplification has been made for (A 3.24).

5. Application of the prediction equations

In this section, results from using the full prediction equations listed in Appendix 3 will be compared to simulation. In the simulations $M = 20$, and F varied from 20 to 200 with two family sizes $n_f = 3$ and 6. For each combination a range of heritabilities from 0 to 0·99 was considered. To assess the predictive power of the method it was necessary to consider three modifications to the methods presented by Wray & Thompson (1990): (i) replacement of binomial sampling by hypergeometric sampling; (ii) correction for the contribution of the base population; and (iii) accounting for the sampling of the breeding value of the ancestors. These modifications are described in Appendices 4–6. The first has greatest influence when $M(F)$ is small, the second when ΔF increases in size, and the third when heritabilities are high. Results are presented in Table 1.

Table 1. *Prediction of ΔF using all terms in Appendix 3 (ΔF_{pred}) compared to rates obtained from simulation (ΔF_{sim}). Values are $100 \times \Delta F$, and % errors are calculated as $(\Delta F_{sim} - \Delta F_{pred})/\Delta F_{pred}$*

		$F = 20$			$F = 40$			$F = 100$			$F = 200$		
	h_0^2	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error
$n_f = 3$	0	104	+3	+2·9	83	0	0	71	0	0	67	−1	−1·5
	0·1	119	+4	+3·4	99	−1	−1·0	88	−6	−6·8	85	−3	−3·5
	0·2	130	+3	+2·3	109	+1	+0·9	98	−4	−4·1	96	−10	−10·4
	0·4	142	0	0	120	−2	−1·6	107	−2	−1·9	105	−5	−4·8
	0·6	145	+5	+3·3	122	+1	+0·8	107	−5	−4·6	104	−5	−4·8
	0·99	137	0	0	109	−1	−0·9	90	−1	−1·1	83	−2	−2·4
$n_f = 6$	0	115	−2	−1·8	89	−1	−1·1	73	0	0	68	0	0
	0·1	146	−2	−1·4	116	+1	+0·9	99	−4	−4·0	95	−8	−8·4
	0·2	166	−5	−3·0	134	−4	−3·0	115	−5	−4·3	111	−8	−7·2
	0·4	187	+5	+2·7	150	0	0	128	−4	−3·1	123	−6	−4·9
	0·6	191	+7	+3·7	152	−2	−1·3	127	−4	−3·1	120	−1	−0·8
	0·99	165	−3	−1·8	124	−5	−4·0	98	−3	−3·1	88	−1	−1·1

* Standard errors from simulations vary from 1 to 2.

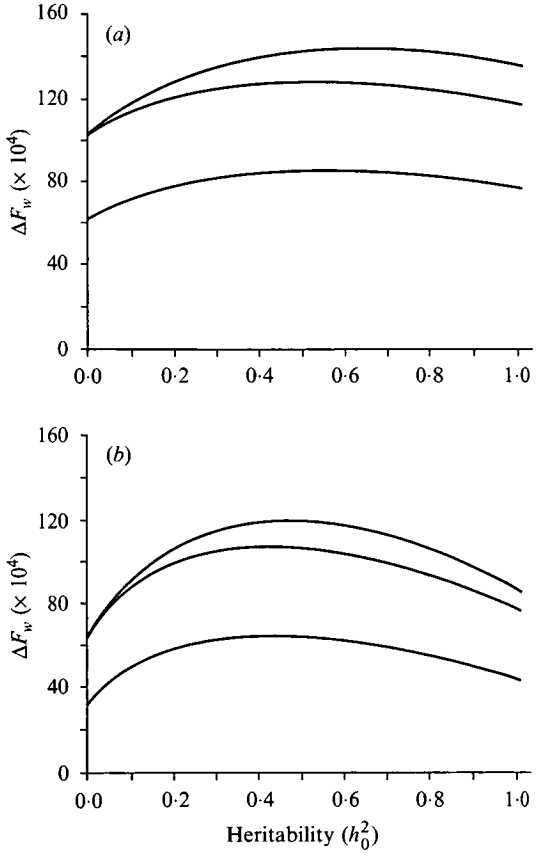


Fig. 1. The relationship of ΔF_w and its three components with heritability for (*a*) $M = F = 20$, $n_f = 3$ and (*b*) $M = 20$, $F = 200$, $n_f = 6$. The lower line is the squared mean contribution, the middle line is the sum of the squared mean and sampling contributions and the upper line is ΔF_w. (ΔF_w is the rate of inbreeding uncorrected for base contributions as described in Appendix 5.)

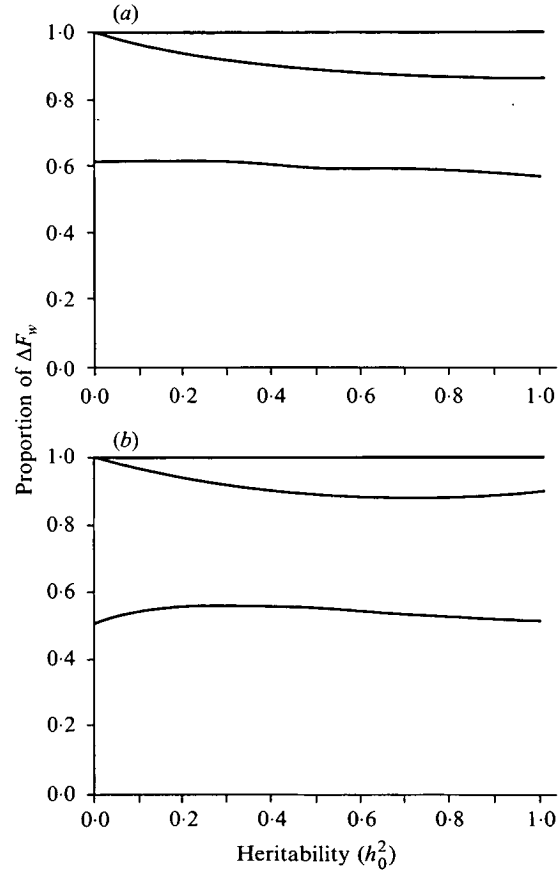


Fig. 2. The proportional contribution of the three components of inbreeding for (*a*) $M = F = 20$, $n_f = 3$ and (*b*) $M = 20$, $F = 200$, $n_f = 6$. The lower line is the contribution of the squared mean and the middle line is the sum of the squared mean and sampling contribution.

The results show a clear trend; very accurate prediction for $F \leqslant 40$, but an increasing tendency to overpredict as F increases further. This is also confounded with the increase in FM^{-1}. The possible reasons will be discussed in more detail in a later section.

It has been shown that for mass selection 3 types of contribution to long-term contributions (and hence to ΔF) can be identified and modelled: squared mean,

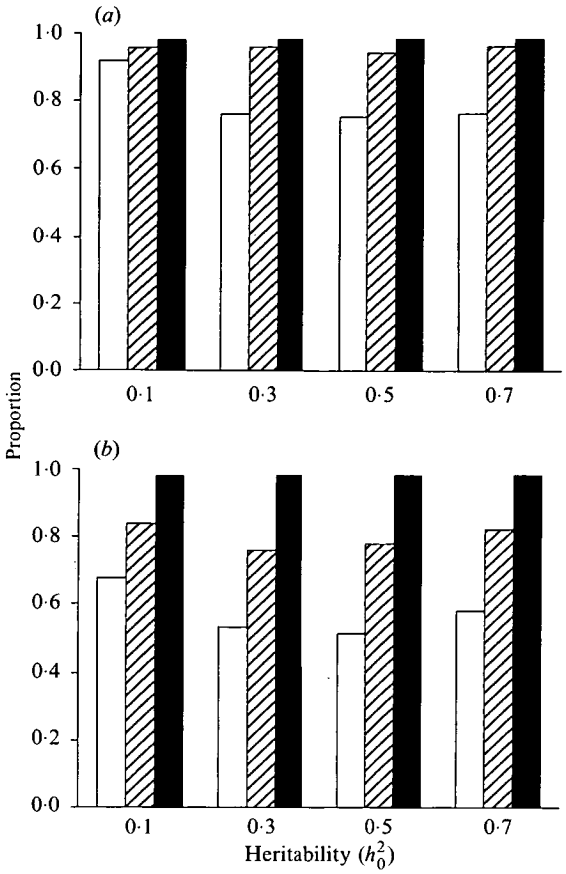


Fig. 3. The proportional contribution of (i) equation (A 3.1) to the squared mean contribution; (ii) equation (A 3.6) to the sampling contribution; and (iii) equations (A 3.13), (A 3.16) and (A 3.17) to the co-selection contribution. Results are shown as (*a*) $M = F = 20$, $n_f = 3$ and (*b*) $M = 20$, $F = 200$, $n_f = 6$. (See Appendix 3 for definition of the equations.)

sampling and co-selection of sibs. These sources change in importance as the population parameters change and further, within each type the importance of particular contributions involving inheritance also varies. Among the terms for the squared mean only (A 3.1) occurs independently of selection, and likewise (A 3.6) is the only sampling term to occur in the absence of selection. For the purpose of this paper, the co-selection of sibs occurs through genetic covariance only. However, terms analogous to (A 3.13), (A 3.16) and (A 3.17) will occur when common family variance is encountered through non-genetic means e.g. maternal or environmental factors, but none of the remaining co-selection terms would occur without some mode of inheritance for the trait under selection. Thus for co-selection there is a case for separating the contribution of these three from the remaining terms among (A 3.13) to (A 3.24).

The magnitude of these contributions have been examined using the prediction for two cases $M = F = 20$, $n_f = 3$ and $M = 20$, $F = 200$ $n_f = 6$, and results are presented in Figs 1–3.

Figure 1 shows the change in the expected long-term squared contribution with heritability separated into three types of contribution, whilst Fig. 2 shows these same data when expressed as a proportion of the total. In the cases considered the predicted $E(r^2)$ reaches its peak for h^2 between 0·4 and 0·6. When $F = 20$, the squared mean contribution remained between 57 and 62% whilst for $F = 200$ the squared mean contribution remained less than 58% of the total. In both examples the proportional contribution from co-selection increased from $h^2 = 0$ to 0·9 and contributed up to 13% of the total.

Figure 3 shows the proportional contribution made by : (i) (A 3.1) to the total squared mean contribution; (ii) (A 3.6) to the sampling contribution; and (iii) (A 3.13), (A 3.16) and (A 3.17) to the co-selection contribution. For h^2 as low as 0·2, approximately 30% of the squared mean contribution was due to inherited advantage, whereas this accounted for up to 20% and only 2% of the sampling and co-selection components respectively. It was also noted that (A 3.10) was large only when $F = 100$ and 200, where ΔF was over-predicted.

6. Approximation and relationship with variance of family size

The following simplifying approximation was strongly suggested by the results.

(i) The terms of the squared mean were all included and when summed give a total contribution of

$$(16M)^{-1}[1+(K+i^2)\rho_m]+(16F)^{-1}[1+(K+i^2)\rho_f], \quad (1)$$

where $K = i^2(S_\infty^2 - 1) + 2iqQ_\infty$ (where S_∞, Q_∞ and q are as defined in Appendix 2, namely $(1-c^{-1})$, $(1-c^2)^{-1}$ and $\frac{1}{2}(i_m c_m + i_f c_f) + ic^2 S_\infty$ respectively).

(ii) Only (A 3.6) was included from the sampling terms and this has the form

$$(16M)^{-1}+(16F)^{-1}-(8T)^{-1}. \quad (2)$$

Correction for hypergeometric sampling is achieved by multiplication with $[\frac{1}{2}(1-M^{-1})+\frac{1}{2}(1-F^{-1})]$.

(iii) Only terms (A 3.13), (A 3.16) and (A 3.17) were included from the covariance terms and these combine to give

$$i^2(16F)^{-1}[\rho_m(FM^{-1}-1)+3(\rho_m+\rho_f)]. \quad (3)$$

More precisely i^2 should be replaced by

$$\tfrac{1}{4}((1-M^{-1})(1-n_f^{-1})i_m^2 + 2i_m i_f + (1-F^{-1})(1-n_f^{-1})i_f^2)$$

although no correction for n_f is necessary for A3.17.

The results from this formula are given in Table 2. Results from Table 2 show that the approximation has a tendency to underpredict for $h^2 = 0{\cdot}4$, 0·6, but exhibits little loss of accuracy for $F \leqslant 40$ but a gain in accuracy for $F = 100$ and 200. It appears more robust.

Latter (1959) and Hill (1972) derived an expression for the rate of inbreeding using the variances of family

Table 2. *Prediction of ΔF using the approximation compared to rates obtained from simulation. Values are $100 \times \Delta F$, and % errors are calculated as $(\Delta F_{sim} - \Delta F_{pred})/\Delta F_{pred}$*

		$F = 20$			$F = 40$			$F = 100$			$F = 200$		
	h_0^2	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$*	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error	ΔF_{pred}	$\Delta F_{sim} - \Delta F_{pred}$	% error
$n_f = 3$	0	104	+3	+2·9	83	0	0	71	0	0	67	−1	−1·5
	0·1	118	+5	+4·2	97	+1	1·0	84	−2	−2·3	81	+1	+1·2
	0·2	128	+5	+3·9	106	+4	+3·8	93	+1	+1·1	90	−4	−4·4
	0·4	140	+2	+1·4	116	+2	+1·8	102	+3	+2·9	98	+2	+2·0
	0·6	144	+6	+4·2	119	+4	+3·4	103	−1	−1	98	+1	+1·0
	0·99	138	−1	+0·7	109	−1	−0·9	89	0	0	82	−1	−1·2
$n_f = 6$	0	115	−2	−1·7	89	−1	−1·1	73	0	0	68	0	0
	0·1	141	+3	+2·1	111	+6	+5·4	94	+1	+1·1	88	−1	−1·1
	0·2	159	+2	+1·3	126	+4	+3·1	107	+3	+2·8	101	+2	+2·0
	0·4	180	+12	+6·6	142	+8	+5·6	119	+5	+4·2	112	+5	+4·5
	0·6	186	+12	+6·1	146	+4	+2·7	120	+3	+2·5	112	+7	+6·2
	0·99	164	−2	+1·2	124	−5	−4·0	97	−2	−2·1	87	0	0

* Standard errors of mean simulated values range from 1 to 2.

size that was general enough to include environmental covariances between sibs but did not include any framework for selection and the inheritance of selective advantage. Wray (1989) derived a form of this equation (ΔF_{LH}) for populations of the same structure considered here that predicted the increased family size arising from sib covariances due to genetic variation and selection in one generation only. This form, ignoring correction factors of $(1-n_f^{-1})$, $(1-M^{-1})$ and $(1-F^{-1})$ is:

$$\Delta F_{LH} = (8M)^{-1} + (8F)^{-1} - (8T)^{-1} + i^2\rho_m[(8M)^{-1} + (8F)^{-1}] + i^2\rho_f(4F)^{-1}$$

This was found to underestimate rates of inbreeding (Wray, 1989; Wray *et al.* 1990). However, if the sum of equations (1) to (3) are denoted ΔF_W [ignoring corrections to (2) and (3)] it can be seen that

$$\Delta F_W = \Delta F_{LH} + K[\rho_m(16M)^{-1} + \rho_f(16F)^{-1}]. \quad (4)$$

In fact, this relationship still holds when the corrections using $(1-n_f^{-1})$, $(1-M^{-1})$ and $(1-F^{-1})$ are made to both ΔF_{LH} and ΔF_W. The term in ΔF_W that is not included in ΔF_{LH}, is part of the squared mean contribution and describes the extra inbreeding arising from selection; this is caused by the interdependence of generations through the inheritance of selected advantage and the consequent expected proliferation of lines arising from superior ancestors in generation 1 at the expense of their inferior contemporaries. Precise prediction will require correction using Appendix 5.

In conclusion, the prediction of inbreeding in mass selection can be shown to approximate closely a prediction involving the variance of family size assuming independent selection processes in each generation plus a single term that describes the cumulative effect of the expected proliferation of lines from superior ancestors in generation 1 at the expense of lines from the inferior ancestors.

(i) *Example*

From $M = 20$, $F = 40$, $n_f = 6$ and $h_0^2 = 0{\cdot}4$: the proportions selected are $p_m = 0{\cdot}0833$ and $p_f = 0{\cdot}1667$; the intensities of selection and variance reduction coefficients are (from standard tables) $i_m = 1{\cdot}839$, $i_f = 1{\cdot}499$ and $i = \frac{1}{2}(i_m + i_f) = 1{\cdot}669$, $k_m = 0{\cdot}838$, $k_f = 0{\cdot}797$ and $k = \frac{1}{2}(k_m + k_f) = 0{\cdot}818$. As is the case for estimating progress, accurate estimation of ΔF requires the calculation of h_2^2 and half-sib correlations (ρ_m and ρ_f). For the example chosen these are $h_2^2 = 0{\cdot}358$, $\rho_m = 0{\cdot}071$ and $\rho_f = 0{\cdot}073$, to give $c_m = 0{\cdot}350$, $c_f = 0{\cdot}357$ and $c = \frac{1}{2}(c_m + c_f) = 0{\cdot}354$, and $K = 7{\cdot}361$.

For simplicity the hypergeometric corrections will be ignored: the squared mean contribution is 0·0081 [equation (1)], the sampling contribution is 0·0042 [equation (2)], and the co-selection contribution is 0·0022 [equation (3)]. This gives a value of $\Delta F_w = 0{\cdot}0145$, which when corrected using Appendix 5 gives $\Delta F = 0{\cdot}0149$ (compared to the value of 0·0150 from stimulation). The prediction is slightly higher than that shown in Table 2 for two reasons: firstly, hypergeometric corrections have not applied; and secondly, the values of h_2^2, ρ_m and ρ_f that were calculated above were not corrected for sampling as described in Appendix 6, but were simply adjusted for the Bulmer effect as defined earlier in the text.

7. Discussion

The work presented has shown that good predictions of rates of inbreeding in mass selection can be made using a straightforward closed expression involving predictable genetic parameters. These genetic parameters involve the intensity of selection and variance

truncation selection with fixed proportion p_x and a correlation ρ_{FS} between the indices of full-sibs, approximation to the required probability is given by $(n_f-1)(T-1)^{-1}$ Prob $(i,j \geqslant v \mid i,j$ full-sibs$) p_x^{-2}$, where v is the truncation deviate for infinite populations. Prob $(i,j \geqslant v \mid i,j$ full-sibs) can be re-expressed as $p_x Q_{FS}$ where $Q_{FS} =$ Prob $(i,j \geqslant v \mid j \geqslant v$ and i,j full-sibs). Thus the required probability is approximately that for $\rho_{FS} = 0$ scaled by the ratio of the conditional and unconditional probabilities of $i \geqslant v$. Mendell & Elston (1974) show $Q_{FS} \approx \Phi[(i_x \rho_{FS} - v)(1-k_x \rho_{FS})^{-\frac{1}{2}}]$ to be an accurate approximation. Thus

Prob $(i,j$ full-sibs $\mid i,j$ selected)

$$\approx (n_f-1)(T-1)^{-1} Q_{FS} p_x^{-1} \approx (n_f-1) Q_{FS} X^{-1}.$$

Similarly

Prob $(i,j$ half-sibs $\mid i,j$ selected)

$$\approx (FM^{-1}-1) n_f Q_{HS} X^{-1}$$

Since individuals are in generation 1, $\rho_{FS} = \frac{1}{2}h_0^2$ and $\rho_{HS} = \frac{1}{4}h_0^2$.

References

Bulmer, M. G. (1971). The effect of selection on genetic variability. *American Naturalist* **105**, 201–221.

Burrows, P. M. (1984*a*). Inbreeding under selection for unrelated families. *Biometrics* **40**, 357–366.

Burrows, P. M. (1984*b*). Inbreeding under selection from related families. *Biometrics* **40**, 895–906.

Hill, W. G. (1972). Effective population size of populations with overlapping generations. *Theoretical Population Biology* **3**, 278–289.

Latter, B. D. H. (1959). Genetic sampling in a random mating population of constant size and sex ratio. *Australian Journal of Biological Science* **12**, 500–505.

Lindgren, D. (1991). Optimal utilization of genetic resources. *Forest Tree Improvement* **23**, 49–67.

Mendell, N. C. & Elston, R. C. (1974). Multifactorial traits: genetic analysis and prediction of recurrence risks. *Biometrics* **30**, 41–57.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.

Tallis, G. M. (1961). The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society* B **23**, 223–229.

Toro, M. A. & Neita, B. (1984). A simple method for increasing the response to artificial selection. *Genetical Research* **44**, 347–349.

Woolliams, J. A. (1989). Modification to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**, 1–14.

Wray, N. R. (1989). Consequences of selection in finite populations with particular reference to closed nucleus herds of swine. Ph.D. Thesis, Edinburgh.

Wray, N. R. & Thompson, R. (1990). Prediction of rates of inbreeding in selected populations. *Genetical Research* **55**, 41–54.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1990). Methods for predicting rates of inbreeding in selected populations. *Theoretical Applied Genetics* **80**, 503–512.

Verrier, E. (1989). Prédiction de l'évolution de la variance genétique dans les populations animales d'effectly limité soumises à selection. Ph.D. Thesis. INA Paris, Grignon, France.

parameters from the second generation of offspring, i.e. heritability, competitiveness ($c = \frac{1}{2}(1-kh^2)$) and half-sib correlations.

The terms involved predict that to order X^{-1}, the rate of inbreeding in the absence of genetic variation and environmental covariation takes the form $(8M)^{-1}+(8F)^{-1}-(8T)^{-1}$ a term that can be derived for random selection with fixed family sizes by other methods (Burrows, 1984*a*; Wray, Woolliams and Thompson, 1990). This paper has shown that when viewed from the concept of long-term contributions this has two components, firstly a squared mean contribution of $(16M)^{-1}+(16F)^{-1}$ and secondly a smaller sampling contribution of $(16M)^{-1}+(16F)^{-1}-(8T)^{-1}$. When heritable and non-heritable correlations are present between sibs in a hierarchical scheme then additional terms analogous to the situations considered by Latter (1959) and Hill (1972) are required plus the addition to the squared mean contribution of $K[\rho_m(16M)^{-1}+\rho_f(16F)^{-1}]$ were ρ_x is the half-sib correlation with common parent of sex x, and $K = i^2(S_\infty^2-1)+2iqQ_\infty$. If selection is assumed to be of equal intensity in the sexes $q = icS_\infty$ and $K = i^2(S_\infty-1)(1+S_\infty+2Q_\infty)$. This extra term describes the inbreeding arising from the expected proliferation of lines arising from superior ancestors in generation 1 at the expense of lines from their inferior contemporaries.

The full derivation of the theory presented by Wray & Thompson (1990) was found to be very accurate in the situations studied in this paper for $F \leqslant 40$, but clearly overestimated for $F \geqslant 100$ where male half-sib families are large. Overestimation can be firmly ascribed to the inability to predict $E(r^2)$ rather than a failure in the relationship of $E(r^2)$ with ΔF: simulations show that for mass selection, index selection and niche selection with environmental covariances, $E(r^2)$ accurately predicts ΔF when mating is at random (N. R. Wray, unpublished results). Two possible reasons for the inability to predict $E(r^2)$ in these circumstances can be advanced. Firstly when modelling the sib covariance only partial adjustment has been made for hypergeometric sampling; when covariances are added, negative covariances must also be added elsewhere i.e. one families success in another's failure. When the half-sib family size is large, selection becomes more intense ($p_m \leqslant n_m^{-1}$) and Appendix 1 shows the regression of the selection score of an ancestor on its breeding value becomes steeper: in the next generation, if an ancestor has many descendants the success of one branch of his family is to the detriment of another. Thus the expected increase in pathways is not as great as predicted. A second contributing cause is that in later generations the accumulation of variances has been assumed to depend on the ancestor [equation (29) of Wray & Thompson, 1990]. In fact after generation 2 families of descendants are a mix of ancestors from the same sex and whilst the overall expected gain in pathways is unaffected by this, the variance of additional contributions between like-sex ancestors will be reduced. One implication of all these considerations is that the robustness of estimation may be less determined by the number of parents than by the number of parents in relation to family size. Nevertheless, the approximation derived is robust over a wide range of parameters, and although possibly benefitting from compensating errors, achieves a great gain in simplicity.

The importance of identifying the components to long-term contributions and quantifying their magnitude in terms of predictable genetic parameters, is that it is only by these means that the value of quanta of information can be assessed for both promoting genetic gain and promoting inbreeding. With this understanding it would then be possible to reconsider selection indices and scheme design to maximise genetic gain while simultaneously constraining inbreeding. The impact of this would clearly depend on the circumstances and chosen constraints. An example of this was given by Woolliams (1989) in which the change from hierarchical to factorial mating designs increased progress when inbreeding was constrained: the foregoing analysis shows that this change left the squared mean and sampling contributions unchanged but substantially reduced that from co-selection. The full potential will, however, require the extension of the foregoing analysis to index selection.

A further finding of the paper is that the expected long-term contribution of an ancestor in mass selection relative to that for random selection will increase by a factor that is linearly related to its breeding value. The factor is of the form $(1+i(1-c)^{-1}A'_{i(x)})$, where $A'_{i(x)}$ is $\frac{1}{2}A_{i(x)}\sigma_P^{-1}$, and $A_{i(x)}$ is the ancestors breeding value about the mean of those selected in generation 1. Thus selection will be expected to leave the contribution unchanged only for an average ancestor, and in general the expected contribution will be linearly related to the ancestors breeding value even though in their offspring in generation 2, contributions are equal. Toro & Neita (1984) and Lindgren (1991) have shown that over a single generation when a restriction is placed on diversity, genetic gain is increased by the differential use of individuals, linearly related to their breeding value. An interesting question arises as to whether over many generations this reduces or exacerbates inbreeding since the superior individuals not only have more progeny but their progeny have a selective advantage. This will determine if linear deployment is a strategy for long-term selection or for clonal propagation.

In a wider context it has been shown that a tractable approach to the problem of predicting inbreeding in the presence of some form of selection may be to estimate the expected long-term squared contribution of ancestors. However, these have only been shown to be equivalent for random mating. It would seem a logical and sensible alternative in some circumstances to assess breeding schemes by $\frac{1}{4}(M+F)^{-2}\Sigma E(r_i^2)$ rather

than by ΔF and that this assessment may also prove tractable in the absence of random mating.

J. A. Woolliams and R. Thompson gratefully acknowledge funding by the Ministry of Agriculture, Fisheries and Food, UK.

Appendix 1: Derivation of $\boldsymbol{b}_{x,t}$

$$\mathbb{D}\boldsymbol{b}_{x,t} = \mathbb{D}\boldsymbol{b}_{x,t-1} + \tfrac{1}{4}MX^{-1}n_m c_m c^{t-3}\sigma_P^{-1}\boldsymbol{z} + \tfrac{1}{4}FX^{-1}n_f c_f c^{t-3}\sigma_P^{-1}\boldsymbol{z}$$
$$= \mathbb{D}\boldsymbol{b}_{x,t-1} + \tfrac{1}{2}n_x c^{t-2}\sigma_P^{-1}\boldsymbol{z}.$$

Since

$$\mathbb{D}^i = \mathbb{D}, \boldsymbol{b}_{x,t} = \mathbb{D}\left[\boldsymbol{b}_{x,2} + \tfrac{1}{2}n_x\sigma_P^{-1}\left(\sum_{i=1}^{t-3} c^i\right)\boldsymbol{z}\right] + \tfrac{1}{2}n_x c^{t-2}\sigma_P^{-1}\boldsymbol{z}.$$

Substituting $\boldsymbol{b}_{x,2} = \tfrac{1}{2}n_x\sigma_P^{-1}\boldsymbol{z}$, $n_x = TX^{-1}$ and $z_x = Xi_x T^{-1}$ and defining $S_t = \Sigma_{i=0}^{t} c^i$, gives

$$\boldsymbol{b}_{x,t} = \tfrac{1}{2}X^{-1}\sigma_P^{-1}[S_{t-3}\boldsymbol{D} + c^{t-2}\boldsymbol{I}](Mi_m, Fi_f)^T.$$

Finally $\boldsymbol{D}(Mi_m, Fi_f)^T = i(M,F)^T$, thus

$$\boldsymbol{b}_{x,t} = \tfrac{1}{2}\sigma_P^{-1}[S_{t-3}\, i(MX^{-1}, FX^{-1})^T + c^{t-2}(i_m MX^{-1}, i_f FX^{-1})].$$

Appendix 2

$$\boldsymbol{P}_{i(x),t} = \prod_{j=t}^{\infty} \boldsymbol{M}_{i(x),j} = \boldsymbol{P}_{i(x),t+1}\cdot\boldsymbol{M}_{i(x),t}$$
$$= \boldsymbol{P}_{i(x),t+1}[\boldsymbol{D} + c^{t-3}A'_{i(x)}\boldsymbol{LDR}],$$

where $\boldsymbol{L}$ and $\boldsymbol{R}$ are diagonal matrices with non-zero elements i_m and i_f for $\boldsymbol{L}$ and c_m and c_f for $\boldsymbol{R}$. Thus

$$\boldsymbol{P}_{i(x),t} = \boldsymbol{P}_{i(x),t+2}[\boldsymbol{D} + c^{t-2}A'_{i(x)}\boldsymbol{LDR}][\boldsymbol{D} + c^{t-3}A'_{i(x)}\boldsymbol{LDR}]$$
$$= \boldsymbol{P}_{i(x),t+2}[\boldsymbol{D} + c^{t-2}A'_{i(x)}\boldsymbol{LDRD} + c^{t-3}A'_{i(x)}\boldsymbol{DLDR} + c^{t-2}c^{t-3}A'^2_{i(x)}\boldsymbol{LDRLDR}].$$

These terms can be simplified by noting $\boldsymbol{DLD} = i\boldsymbol{D}$ and $\boldsymbol{DRD} = c\boldsymbol{D}$, and likewise for other diagonal matrices; and, if terms of order $A^3_{i(x)}$ or higher are ignored, by similar multiplication and collection of terms

$$\boldsymbol{P}_{i(x),t} = \boldsymbol{D} + A'_{i(x)}\left[\left(\sum_{t-2}^{\infty} c^j\right)ic\boldsymbol{D} + c^{t-3}i\boldsymbol{DR}\right] + A'^2_{i(x)}\left[\left(\sum_{t-2}^{\infty} c^j c^{j+1}\right)icg\boldsymbol{D} + \left(\sum_{j\geqslant t-2,\, k>j+1}^{\infty} c^j c^k\right)i^2c^2\boldsymbol{D} + c^{t-3}c^{t-2}ig\boldsymbol{DR} + c^{t-3}\sum_{t-1}^{\infty} c^j i^2 c\boldsymbol{DR}\right].$$

where $g_x = i_x c_x$ and $g = \tfrac{1}{2}(g_m + g_f)$. If

$$S_\infty = \sum_{j=0}^{\infty} c^j = (1-c)^{-1}, \quad \text{and} \quad Q_\infty = \sum_{j=0}^{\infty}(c^j)^2 = (1-c^2)^{-1}$$

$$\boldsymbol{P}_{i(x),t} = \boldsymbol{D} + A'_{i(x)}ic^{t-3}\boldsymbol{D}[c^2S_\infty\boldsymbol{I} + \boldsymbol{R}] + A'^2_{i(x)}ic^{2t-5}[g + ic^2S_\infty]\,\boldsymbol{D}[c^3Q_\infty\boldsymbol{I} + \boldsymbol{R}].$$

This can be further simplified by defining terms $q = g + c^2iS_\infty$, and diagonal matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Lambda}$ with elements θ_m, θ_f and λ_m, λ_f where $\theta_w = c_w + c^2S_\infty$ and $\lambda_w = c_w + c^3Q_\infty$, then

$$\boldsymbol{P}_{i(x),t} = \boldsymbol{D} + A'_{i(x)}ic^{t-3}\boldsymbol{D\Theta} + A'^2_{i(x)}ic^{2t-5}q\boldsymbol{D\Lambda}.$$

Note

$$\tfrac{1}{2}(\theta_m + \theta_f) = cS_\infty \quad \text{and} \quad \tfrac{1}{2}(\lambda_m + \lambda_f) = cQ_\infty.$$

Appendix 3

Listing of terms for ΔF derived for sex x. The constants used in terms are

$$B_\infty = \sum_{j=0}^{\infty} c^{2j}2^{-j} \quad \text{and} \quad H_\infty = \sum_{j=0}^{\infty} c^j 2^{-j}.$$

Note

$$\sum_{j=0}^{\infty} c^j S_j 2^{-j} = B_\infty H_\infty.$$

(i) *Squared mean contribution*

v_0: $(16X)^{-1}$. (A 3.1)

v_1: $\rho_x i^2(16X)^{-1}$. (A 3.2)

v_2: $\rho_x i^2c^2S_\infty^2(16X)^{-1}$. (A 3.3)

v_3: $2\rho_x(i^2cS_\infty + iq)(16X)^{-1}$. (A 3.4)

v_4: $2\rho_x ic^2qQ_\infty(16X)^{-1}$. (A 3.5)

(ii) *Binomial sampling*

Terms arising from V_0

v_0: $(M^{-1} + F^{-1} - 2T^{-1})/32$. (A 3.6)

v_2: $\rho_x B_\infty i^2(\theta_m^2 M^{-1} + \theta_f^2 F^{-1} - (\theta_m^2 + \theta_f^2)T^{-1})/64$. (A 3.7)

v_4: $\rho_x B_\infty icq(\lambda_m M^{-1} + \lambda_f F^{-1} - 2\lambda T^{-1})/32$. (A 3.8)

Terms arising from $A'_{i(x)}V_0$

v_3: $\rho_x B_\infty H_\infty i^2c(\theta_m M^{-1} + \theta_f F^{-1} - 2\theta T^{-1})/64$. (A 3.9)

Terms arising from $A'_{i(x)}V_1$

v_3: $\rho_x B_\infty i(\theta_m i_m M^{-1} + \theta_f i_f F^{-1} - 4qT^{-1})/32$. (A 3.10)

Terms arising from $A'^2_{i(x)}V_1$

$v_1 = \rho_x B_\infty(g + \tfrac{1}{2}ic^2H_\infty)(i_m M^{-1} + i_f F^{-1} - 4iT^{-1})/128$. (A 3.11)

Terms arising from $A'^2_{i(x)}V_2$

$v_1 = -\rho_x(1 + \tfrac{1}{4}B_\infty(c_m^2 + c_f^2))(i_m^2 + i_f^2)T^{-1}/64$. (A 3.12)

(iii) *Co-selection of full-sibs in generation 2*

Terms arising from C_0

v_0: $\rho_{x'} i^2(16F)^{-1}$. (A 3.13)

$v_2 = \rho_x\rho_{x'} i^2q^2(16F)^{-1}$. (A 3.14)

$v_4 = \rho_x\rho_{x'} i^2cq(\lambda_m i_m + \lambda_f i_f)(16F)^{-1}$. (A 3.15)

(iv) *Co-selection of sibs in generation $t \geqslant 3$*

Terms arising from $\boldsymbol{C}_0$

v_0: $(\rho_m+\rho_f)\,i^2(16F)^{-1}$. (A 3.16)

v_0: $\rho_m(d-1)\,i^2(32F)^{-1}$. (A 3.17)

v_0: $-B_\infty i^2(\rho_m c_m^2 M^{-1}+\rho_f c_f^2 F^{-1})/64$. (A 3.18)

v_2: $\rho_x(\rho_m+\rho_f)\,B_\infty i^2c^2q^2(32F)^{-1}$. (A 3.19)

v_2: $\rho_x\rho_m(d-1)\,B_\infty i^2c^2q^2(64F)^{-1}$. (A 3.20)

v_4: $\rho_x(\rho_m+\rho_f)\,B_\infty i^2c^3q(\lambda_m i_m+\lambda_f i_f)(32F)^{-1}$. (A 3.21)

v_4: $\rho_x\rho_m(d-1)\,B_\infty i^2c^3q(\lambda_m i_m+\lambda_f i_f)(64F)^{-1}$. (A 3.22)

Terms arising from $A'_{i(x)}\boldsymbol{C}_0$

v_3: $\rho_x(\rho_m+\rho_f)\,B_\infty H_\infty i^3cq(16F)^{-1}$. (A 3.23)

v_3: $\rho_x\rho_m(d-1)\,B_\infty H_\infty i^3cq(32F)^{-1}$. (A 3.24)

Appendix 4. Hypergeometric sampling

For sampling without replacement the variance of the hypergeometric distribution is more appropriate than that of the binomial. When selecting W individuals at random from a total T without replacement the variance of a total family contribution for a family of $n_w = TW^{-1}$ members is $[1-(n_w-1)(T-1)^{-1}]$ times the binomial variance. When terms in T^{-1} are neglected this factor is approximately $(1-W^{-1})$ and the variance is then $(1-W^{-1})\,n_w p(1-p)$ where $p = WT^{-1}$. In derivation p is a function of family and sex but these deviations are assumed to have no effect on the factor $(1-W^{-1})$. Terms used to derive (A 3.6) to (A 3.12) inclusive require these corrections. In the derivation of terms arising from the co-selection of sibs the simplifying assumption was made that the joint probability of selecting two sibs of like sex is p^2 where again p is a function of the sex and family. This probability is better approximated by $p^2(1-W^{-1})$ and the correction factor $(1-W^{-1})$ is used multiplicatively with the correction $(1-n_w^{-1})$ which is described in the text. Since equations (A 3.13) to (A 3.24) include both like-sex and unlike-sex contributions these correction terms are not simple factors.

Appendix 5. Contributions from the base population

In their derivation Wray & Thompson (1990) show $\Delta F \approx C_1(2-C_0)^{-1}$ where C_1 and C_0 are the average of the diagonal elements in the genetic contribution matrices for generations 1 and 0 (i.e. the unselected base) respectively. They argued that since C_0 is small, $\Delta F \approx \frac{1}{2}C_1$. However, C_0 can contribute to significant proportional errors even when ΔF is in the range covered by the mass selection examples in their paper. However, by using the relations $\Delta F \approx \frac{1}{4}C_0$ (J. A. Woolliams, N. R. Wray, unpublished results) this problem may be largely overcome and the need for predicting C_0 avoided.

$\Delta F = C_1(2-C_0)^{-1} = C_1(2-4\Delta F)^{-1}$ since $\Delta F \approx \frac{1}{4}C_0$ and expanding to a quadratic gives $4\Delta F^2-2\Delta F+C_1=0$. Solving for the lowest root, $\Delta F = \frac{1}{4}(1-(1-4C_1)^{\frac{1}{2}})$.

Substituting ΔF_w for $\frac{1}{2}C_1$, the estimate used by Wray & Thompson (1990), gives

$$\Delta F = \tfrac{1}{4}(1-(1-8\Delta F_w)^{\frac{1}{2}}) \approx \Delta F_w(1+2\Delta F_w),$$

the latter approximation arising from expanding $(1-8\Delta F_w)^{\frac{1}{2}}$ as Taylor series. Thus when $\Delta F \approx 0{\cdot}05$, ΔF_w underestimates ΔF by 10%.

Appendix 6. Sampling of ancestral breeding values

The average contribution from sex x to sex y, $r_{i(x)y\infty}$, is constrained to be $\frac{1}{2}YX^{-1}$. The sampling distributions incorporated ensure the error terms obey this constraint. However, no allowance has so far been made for $\mu_{i(x)y,\infty}$; as modelled the average expected contribution is

$$X^{-1}\sum_{i=1}^{X}\mu_{i(x)y,\infty} = \tfrac{1}{2}YX^{-1}(1+iS_\infty \bar{A}'_{i(x)}).$$

The deviation from the constrained value decreases as X increases since $\mathrm{Var}(\bar{A}_{i(x)})$ becomes smaller and $\bar{A}'_{i(x)}$ lies more surely close to 0.

Two approaches can be adopted to overcome this problem which are equivalent in first order terms. Firstly, the coefficients $\mu_{i(x)y,t}$ and $\mu_{i(x)wy,t}$, can be recalculated with the discrete distribution of genotypes obtained from selection in generation 1 and, making the same approximations as were made for the truncated normal, it is seen that $A'_{i(x)}$ should be replaced by $A'_{i(x)}-\bar{A}'_{i(x)}$. Alternatively, the $\mu_{i(x)y,t}$ and $\mu_{i(x)wy,t}$ may be regarded as fitness coefficients, and the process can be recalculated using relative fitness to replace absolute fitness, by dividing through by mean fitness coefficients at each generation and transition. The use of $A'_{i(x)}-\bar{A}'_{i(x)}$ requires the estimation of its variance.

$$\mathrm{Var}(A_{i(x)}-\bar{A}_{i(x)}) = \mathrm{Var}(A_{i(x)})-\mathrm{Var}(\bar{A}_{i(x)}),$$

$$\mathrm{Var}(\bar{A}_{i(x)}) = X^{-1}\mathrm{Var}(A_{i(x)})+X^{-2}\sum_{i\neq j}^{X}\mathrm{cov}(A_{i(x)},A_{j(x)})$$

$$= X^{-1}\mathrm{Var}(A_{i(x)})+(1-X^{-1})$$
$$[\mathrm{Prob}(i,j\ \text{half-sibs}\,|\,i,j\ \text{selected})$$
$$\times \mathrm{cov}(\text{half-sibs})$$
$$+\mathrm{Prob}(i,j\ \text{full-sibs}\,|\,i,j\ \text{selected})$$
$$\times \mathrm{cov}(\text{full-sibs})].$$

After selection in generation 1, $\mathrm{Var}(A_{i(x)}) = h_0^2(1-k_x h_0^2)$, $\mathrm{cov}(\text{half-sibs}) \approx \frac{1}{4}h_0^2(1-k_x h_0^2)$ and $\mathrm{cov}(\text{full-sibs}) \approx \frac{1}{2}h_0^2(1-k_x h_0^2)$. More precise estimates of covariance, useful for high $h^2 > 0{\cdot}8$ can be obtained by applying the results of Tallis (1961).

The results of Wray *et al.* (1990) are used to approximate $\mathrm{Prob}(i,j\ \text{full-sibs}\,|\,i,j\ \text{selected})$ when sampling without replacement by standard normal probabilities. For a normally distributed trait undergoing

truncation selection with fixed proportion p_x and a correlation ρ_{FS} between the indices of full-sibs, approximation to the required probability is given by $(n_f-1)(T-1)^{-1}$ Prob $(i,j \geqslant v \mid i,j$ full-sibs$)p_x^{-2}$, where v is the truncation deviate for infinite populations. Prob $(i,j \geqslant v \mid i,j$ full-sibs) can be re-expressed as $p_x Q_{FS}$ where Q_{FS} = Prob $(i,j \geqslant v \mid j \geqslant v$ and i,j full-sibs). Thus the required probability is approximately that for $\rho_{FS} = 0$ scaled by the ratio of the conditional and unconditional probabilities of $i \geqslant v$. Mendell & Elston (1974) show $Q_{FS} \approx \Phi[(i_x \rho_{FS} - v)(1-k_x \rho_{FS})^{-\frac{1}{2}}]$ to be an accurate approximation. Thus

Prob $(i,j$ full-sibs$\mid i,j$ selected)

$$\approx (n_f-1)(T-1)^{-1} Q_{FS} p_x^{-1} \approx (n_f-1) Q_{FS} X^{-1}.$$

Similarly

Prob $(i,j$ half-sibs$\mid i,j$ selected)

$$\approx (FM^{-1}-1) n_f Q_{HS} X^{-1}$$

Since individuals are in generation 1, $\rho_{FS} = \frac{1}{2}h_0^2$ and $\rho_{HS} = \frac{1}{4}h_0^2$.

References

Bulmer, M. G. (1971). The effect of selection on genetic variability. *American Naturalist* **105**, 201–221.

Burrows, P. M. (1984*a*). Inbreeding under selection for unrelated families. *Biometrics* **40**, 357–366.

Burrows, P. M. (1984*b*). Inbreeding under selection from related families. *Biometrics* **40**, 895–906.

Hill, W. G. (1972). Effective population size of populations with overlapping generations. *Theoretical Population Biology* **3**, 278–289.

Latter, B. D. H. (1959). Genetic sampling in a random mating population of constant size and sex ratio. *Australian Journal of Biological Science* **12**, 500–505.

Lindgren, D. (1991). Optimal utilization of genetic resources. *Forest Tree Improvement* **23**, 49–67.

Mendell, N. C. & Elston, R. C. (1974). Multifactorial traits: genetic analysis and prediction of recurrence risks. *Biometrics* **30**, 41–57.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.

Tallis, G. M. (1961). The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society* B **23**, 223–229.

Toro, M. A. & Neita, B. (1984). A simple method for increasing the response to artificial selection. *Genetical Research* **44**, 347–349.

Woolliams, J. A. (1989). Modification to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**, 1–14.

Wray, N. R. (1989). Consequences of selection in finite populations with particular reference to closed nucleus herds of swine. Ph.D. Thesis, Edinburgh.

Wray, N. R. & Thompson, R. (1990). Prediction of rates of inbreeding in selected populations. *Genetical Research* **55**, 41–54.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1990). Methods for predicting rates of inbreeding in selected populations. *Theoretical Applied Genetics* **80**, 503–512.

Verrier, E. (1989). Prédiction de l'évolution de la variance genétique dans les populations animales d'effectly limité soumises à selection. Ph.D. Thesis. INA Paris, Grignon, France.

Paper 11

Theor Appl Genet (1994) 87:878–892

Prediction of rates of inbreeding in populations undergoing index selection

N. R. Wray[1], J. A. Woolliams[2], R. Thompson[2]

[1] Livestock Improvement Unit, Victorian Institute of Animal Science, 475 Mickleham Road, Attwood, Victoria 3049, Australia
[2] AFRC Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, Scotland, UK

Received: 22 February 1993 / Accepted: 16 June 1993

Abstract. For populations undergoing mass selection, previous studies have shown that the rate of inbreeding is directly related to the mean and variance of long-term contributions from ancestors to descendants, and thus prediction of the rate of inbreeding can be achieved via the prediction of long-term contributions. In this paper, it is shown that the same relationship between the rate of inbreeding and long-term contributions is found when selection is based on an index of individual and sib records (index selection) and where sib records may be influenced by a common environment. In these situations, rates of inbreeding may be considerably higher than under mass selection. An expression for the rate of inbreeding is derived for populations undergoing index selection based on variances of (one-generation) family size and incorporating the concept of long-term selective advantage. When the mating structure is hierarchical, and when half-sib records are included in the index, the correlation between parental breeding values and the index values of their offspring is higher for male parents than female parents. This introduces an important asymmetry between the contributions of male and female ancestors to the evolution of inbreeding which is not present when selection is based on individual and/or full-sib records alone. The prediction equation for index selection accounts for this asymmetry. The prediction is compared to rates of inbreeding calculated from simulation. The prediction is good when family size is small relative to the number selected. The reasons for over-prediction in other situations are discussed.

Key words: Inbreeding rate – Effective population size – Index selection

Introduction

The mean and variance of long-term genetic contributions from ancestors (in a closed population) can be related to the rate of inbreeding (Wray and Thompson 1990). After several generations, the long-term contributions from an ancestor stabilise and are the same to all individuals born into the population, with the values differing between ancestors. The mean simply reflects the constraint of the number of ancestors and the parents used in each generation and is the same whether or not selection is practised. The variance, however, is increased by selection.

Offspring of parents who are genetically superior for the trait under selection are more likely to be selected than the offspring of genetically-average or inferior parents. The parents are said to confer a selective advantage to their offspring. When parents are selected at random, the sampling of parents is independent of the sampling process of the previous generation. However, when parents are selected on a heritable trait, the selective advantage is inherited, and is conveyed, in part, from parent to offspring. Thus the breeding value of an ancestor has influences on selection decisions in all subsequent generations. Wray et al. (1990) introduced the terminology of one-generation, two-generation and long-term selective advantage, referring to that conveyed from parent to offspring, grandparent to grandoffspring, and ancestor to (distant) descendant. They reviewed different methods to predict the rate of inbreeding which can be

Communicated by J. S. F. Barker
Correspondence to: N. R. Wray

classified in the same way, according to the number of generations of selective advantage they attempt to incorporate. Only the methods of Robertson (1961), Wray and Thompson (1990) and Woolliams et al. (1993) are long-term methods.

For populations undergoing mass selection Wray and Thompson (1990) presented a recursive algorithm to predict the mean and variance of long-term contributions and hence to predict the rate of inbreeding. Woolliams et al. (1993) modified components of the prediction and presented an explicit expression for the long-term selective advantage and the rate of inbreeding. Further, they showed that the important terms of the prediction can be related back to the equations of Latter (1959) and Hill (1979), with the addition of a term describing the contribution of the expected long-term selective advantage.

The present paper is concerned primarily with the prediction of the rate of inbreeding when selection is based on an index of records of an individual and its collateral relatives and where sib records may be influenced by a common environment. In these situations, the rate of inbreeding may be considerably higher than under mass selection. Justified by the formal derivation of Woolliams et al. (1993), this paper presents a more intuitive derivation which highlights more explicitly the relationship between the long-term contribution method of Wray and Thompson (1990) with that of Robertson (1961), on the one hand, and Latter (1959) and Hill (1979), on the other.

Methods

Definition of population structure and index parameters

Throughout, conventions on notation follow as closely as possible those of Woolliams et al., (1993). The population structure considered is one of hierarchical random mating of F females with M males ($M \leq F$) with discrete generations. Generation 1 is produced by the mating structure from an unrelated, unselected base population. The term 'ancestors' is used to refer to individuals born and selected in generation 1. Each female produces a family of n offspring comprising n_f males and n_f females ($n = 2n_f$). Each male has n_m offspring of each sex ($n_m = n_f F/M$). X, Y or W and subscripts x, y or w may be used to specify a single sex, either male or female e.g., $X = M$ or F, $n_x = n_m$ or n_f. T is used to denote the total number of offspring of each sex born each generation, $T = Xn_x$. The proportion selected is p after truncation at the standardised normal deviate v, $p = \Phi(v)$ with the corresponding normal ordinate $z = \phi(v)$, where $\Phi(.)$ and $\phi(.)$ represent the cumulative and probability density functions of the normal distribution. The standardised selection intensity is $i = z/p$ and the variance reduction factor is $k = i(i - v)$. When these terms have subscripts they are the values for the sex of animals indicated by the subscript, and without subscripts they are the average of the sexes.

Assuming an infinitesimal model of gene effects, the total genetic variance of individuals born in generation t is $\sigma^2_{A,t}$, which can be decomposed as,

$$\sigma^2_{A,t} = \sigma^2_{Am,t-1} + \sigma^2_{Af,t-1} + \sigma^2_{Aw}$$

where $\sigma^2_{Am,t-1}$ and $\sigma^2_{Af,t-1}$ are $\frac{1}{4}$ of the genetic variances between sires and dams born and selected in generation $t-1$, σ^2_{Aw} is the within-family genetic variance, and $\sigma^2_{Am,0} = \sigma^2_{Af,0} = \frac{1}{2}\sigma^2_{Aw} = \frac{1}{4}\sigma^2_{A,0}$. These parameters are used only for $t \leq 2$, and so reductions in genetic variance due to inbreeding are ignored. The phenotypic variance in generation t is,

$$\sigma^2_{P,t} = \sigma^2_{A,t} + \sigma^2_C + \sigma^2_E$$

where σ^2_C and σ^2_E are the common environmental variance of full-sibs and the error variance respectively. Heritability in generation t is defined as $h^2_t = \sigma^2_{A,t}/\sigma^2_{P,t}$.

Selection is assumed to be based on an index (I_H) of individual record (P), the mean of n full-sib records (including individual) ($\bar{P}_D$), and the mean of $(F/M)n$ half-sib (including the individual and its full-sibs) records ($\bar{P}_H$),

$$I_H = \beta_1(P - \bar{P}_D) + \beta_2(\bar{P}_D - \bar{P}_H) + \beta_3\bar{P}_H,$$

where β_1, β_2, β_3 are selection index weights. The index is written in this way because $\mathrm{Cov}(P - \bar{P}_D, \bar{P}_D - \bar{P}_H) = \mathrm{Cov}(P - \bar{P}_D, \bar{P}_H) = \mathrm{Cov}(\bar{P}_F - \bar{P}_H, \bar{P}_H) = 0$. P is defined so that the mean of P each generation is zero. Selection index weights are assumed to be constant throughout and are derived so that the index is optimum in the first generation: $\mathrm{Cov}(A, I) = V(I)$, where A is the individual's breeding value. This assumption is adopted for simplicity and for comparison of prediction results with simulation results later; the theory can be developed analogously without this assumption (Woolliams and Wray, in preparation). The selection index weights each generation are,

$$\beta_1 = \frac{\sigma^2_{Aw}}{\sigma^2_{Aw} + \sigma^2_E}, \quad \beta_2 = \frac{\sigma^2_{Af,0} + \frac{1}{n}\sigma^2_{Aw}}{\sigma^2_{Af,0} + \sigma^2_C + \frac{1}{n}(\sigma^2_{Aw} + \sigma^2_E)}$$

and

$$\beta_3 = \frac{\sigma^2_{Am,0} + \frac{M}{F}\left(\sigma^2_{Af,0} + \frac{1}{n}\sigma^2_{Aw}\right)}{\sigma^2_{Am,0} + \frac{M}{F}\left[\sigma^2_{Af,0} + \sigma^2_C + \frac{1}{n}\left(\sigma^2_{Aw} + \sigma^2_E\right)\right]}.$$

Before selection the variance of the indices of individuals born in generation t is,

$$\begin{aligned}\sigma^2_{I,t} = {} & \beta_1^2(\sigma^2_{Aw} + \sigma^2_E)\left(1 - \frac{1}{n}\right) \\ & + \beta_2^2\left[\sigma^2_{Af,t-1} + \sigma^2_C + \frac{1}{n}(\sigma^2_{Aw} + \sigma^2_E)\right]\left(1 - \frac{M}{F}\right) \\ & + \beta_3^2\left\{\sigma^2_{Am,t-1} + \frac{M}{F}\left[\sigma^2_{Af,t-1} + \sigma^2_C + \frac{1}{n}(\sigma^2_{Aw} + \sigma^2_E)\right]\right\}. \qquad (1)\end{aligned}$$

Before selection $\mathrm{Cov}(A \text{ of sex } x \text{ parent}, I_H \text{ of offspring})_t = 2\tau_x\sigma^2_{Ax,t-1}$ for offspring born in generation t, where

$$\tau_m = \beta_3 \quad \text{and} \quad \tau_f = \beta_2(1 - M/F) + \beta_3(M/F). \qquad (2)$$

In the absence of selection $\mathrm{Cov}(I_H \text{ of sex } x \text{ parent}, I_H \text{ of offspring}) = \frac{1}{2}\tau_x\sigma^2_{I,0}$.

The correlations between indices of two full-sibs ($\rho_{D,t}$) and two half-sibs ($\rho_{H,t}$) are

$$\rho_{D,t} = \left\{ -\beta_1^2 \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2) + \beta_2^2\left(1 - \frac{M}{F}\right) \cdot \left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right] + \beta_3^2\left[\sigma_{Am,t-1}^2 + \frac{M}{F}\left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right]\right]\right\}\bar{\sigma}_{I,t}^{-2}, \quad (3)$$

and

$$\rho_{H,t} = \left\{ -\beta_2^2 \frac{M}{F}\left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right] + \beta_3^2\left[\sigma_{Am,t-1}^2 + \frac{M}{F}\left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right]\right]\right\}\bar{\sigma}_{I,t}^{-2}. \quad (4)$$

The correlation between indices of two full-sibs due to the breeding values of the dam is $\rho_{f,t}$ and the correlation between indices of two full-sibs or two half-sibs due to the breeding value of their sire is $\rho_{m,t}$, where

$$\rho_{x,t} = \tau_x^2 \sigma_{Ax,t-1}^2 \bar{\sigma}_{I,t}^{-2}. \quad (5)$$

Two other selection indices are considered: an index of individual record and full-sib mean, $I_D = \beta_1(P - \bar{P}_D) + \beta_2 \bar{P}_D$, where β_1 is the same as in I_H and

$$\beta_2 = \frac{\sigma_{Am,0}^2 + \sigma_{Af,0}^2 + \frac{1}{n}\sigma_{Aw}^2}{\sigma_{Am,0}^2 + \sigma_{Af,0}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)},$$

and an index of individual record only (mass selection), $I_P = \beta_1 P$, where $\beta_1 = h^2$. In the methodology that follows, derivations are made for index I_H. The results are also appropriate for I_D if in I_H, β_3 is set equal to β_2 of I_D and for I_P if in I_H, β_3 and β_2 are set equal to β_1 of I_P. For each index $\bar{\sigma}_{I,t}^2$ must be calculated appropriately.

Simplification of the method of Wray and Thompson (1990)

Rate of inbreeding from long-term contributions. Under the assumption of constant rate of inbreeding each year, Wray and Thompson (1990) presented an expression for the rate of inbreeding (ΔF) appropriate for selected populations,

$$\Delta F \approx \frac{1}{4(M+F)^2} \sum_{j=1}^{M+F} r_{j,\infty}^2 \quad (6)$$

where $r_{j,t}$ is the total additive genetic contribution of ancestor j born in generation 1 to its descendants born in generation t. Alternatively, $r_{j,t}/(M+F)$ is the additive genetic relationship between the Mendelian sampling term that ancestor j received and each descendant, i.e., the genetic relationship between ancestor and descendant which cannot be traced to the base generation (the parents of the ancestors). Appendix 1 shows that this expression can be partitioned by sex of ancestor and by sex of descendant to give,

$$\Delta F \approx \frac{1}{16}\left\{\frac{1}{M^2}\left[\sum_{j=1}^{M} r_{j(mm),\infty}^2 + \sum_{j=1}^{F} r_{j(fm),\infty}^2\right] + \frac{2}{MF}\left[\sum_{j=1}^{M} r_{j(mm),\infty} r_{j(mf),\infty} + \sum_{j=1}^{F} r_{j(fm),\infty} r_{j(ff),\infty}\right] + \frac{1}{F^2}\left[\sum_{j=1}^{M} r_{j(mf),\infty}^2 + \sum_{j=1}^{F} r_{j(ff),\infty}^2\right]\right\} \quad (7)$$

where $r_{j(xy),t}$ is the long-term contribution of ancestor j of sex x to its descendants of sex y. Terms $r_{j(xy),t}$ have mean $\mu_{r(xy),t}$, variance $\sigma_{r(xy),t}^2$ and covariance between male and female descendants of $\sigma_{r(xm,xf),t}$. Therefore, an equivalen expression can be written,

$$E[\Delta F] \approx \frac{1}{16M}\left\{[\mu_{r(mm),\infty}^2 + \sigma_{r(mm),\infty}^2] + 2\left[\frac{M}{F}\right][\mu_{r(mm),\infty}\mu_{r(mf),\infty} + \sigma_{r(mm,mf),\infty}] + \left[\frac{M}{F}\right]^2[\mu_{r(mf),\infty}^2 + \sigma_{r(mf),\infty}^2]\right\} + \frac{1}{16F}\left\{[\mu_{r(ff),\infty}^2 + \sigma_{r(ff),\infty}^2] + 2\left[\frac{F}{M}\right][\mu_{r(fm),\infty}\mu_{r(ff),\infty} + \sigma_{r(fm,ff),\infty}] + \left[\frac{F}{M}\right]^2[\mu_{r(fm),\infty}^2 + \sigma_{r(fm),\infty}^2]\right\}. \quad (8)$$

Mean of long term contributions. The mean $\mu_{r(xy),t} = E[r_{j(xy),t}]$ is an expectation conditional on the deviation of the breeding value of ancestor j over that of its selected contemporaries of $A_{j(x)}$, such that $E[A_{j(x)}] = 0$ and $V[A_{j(x)}] = 4\sigma_{Ax,1}^2$ (which is the variance about the mean of all selected ancestors of sex x, evaluated in Appendix 2). Assuming a linear model, the mean can be expressed as

$$\mu_{r(xy),t} = E[r_{j(xy),t}] = 2^{t-2}E\left[\frac{Y}{X} + b_{xy,t}A_{j(x)}\right]\left(\frac{1}{2}\right)^{t-1} = \frac{1}{2}E\left[\frac{Y}{X} + b_{xy,t}A_{j(x)}\right] \quad (9)$$

(Wray and Thompson 1990). The term $2^{t-2}b_{xy,t}$ can be interpreted as the regression coefficient of the number of distinct pedigree pathways to descendants of sex y in generation t on the breeding values of their ancestors of sex x. The term $(\frac{1}{2})^{t-1}$ represents the relationship between ancestor and descendant along a single pathway. When selection is at random $b_{xy,t}$ is zero. Under selection, $b_{xy,2}$ is the one-generation selective advantage and $b_{xy,\infty}$ is the long-term selective advantage.

Under mass selection Wray and Thompson (1990) showed that,

$$b_{xy,2} \approx \frac{1}{2\sigma_{P,2}} n_x z_y = \frac{1}{2\sigma_{P,2}} \frac{Y}{X} i_y,$$

where $\sigma_{P,2}$ is the phenotypic standard deviation in generation 2. They presented a recursion to calculate $b_{xy,t}$ and ultimately $b_{xy,\infty}$. However, Woolliams et al. (1993) derived a direct expression for $b_{xy,t}$, and showed that,

$$b_{xy,\infty} \approx \frac{1}{2\sigma_{P,2}} \frac{Y}{X} i S_\infty = \frac{1}{\sigma_{P,2}} \frac{Y}{X} \frac{i}{(1 + kh_2^2)}, \quad (10)$$

where S_∞ represents the sum of an infinite series and $S_\infty = (1-c)^{-1}$ with $c = 0.5(1 - kh_2^2)$ defined as the 'coefficient of com-

petitiveness'. For selection indices I_H or I_D (or I_P) the expressions for $b_{xy,2}$ and $b_{xy,\infty}$ are derived in Appendix 2 resulting in,

$$b_{xy,2} \approx \frac{1}{2\sigma_{I,2}} n_x \tau_x z_y = \frac{1}{2\sigma_{I,2}} \frac{Y}{X} \tau_x i_y$$

$$b_{xy,\infty} \approx \frac{1}{2\sigma_{I,2}} \frac{Y}{X} i S_\infty \frac{(\tau + \tau_x)}{2}(1-\psi) = \frac{1}{2\sigma_{I,2}} \frac{Y}{X} i \frac{(\tau + \tau_x)}{1 + k\tau}(1-\psi) \tag{11}$$

where S_∞ is defined as for mass selection as $(1-c)^{-1}$ but the coefficient of competitiveness is more generally defined as $c = \frac{1}{2}(1 - k\tau)$ with $\tau = \frac{1}{2}(\tau_m + \tau_f)$ and τ_x is defined in equation (2) and where $\psi = (k_m - k_f)(\tau_m - \tau_f)/8$. For selection on index I_p (mass selection) where $\tau_m = \tau_f = h^2$ and $\sigma_{I,2} = h^2 \sigma_{P,2}$, equation (11) reduces to equation (10) except that h_2^2 is replaced by $\tau = h^2$ here, where h_2^2 arose from a more-accurate approximation to equation (A2.2) than the one used here for reasons of complexity with index selection.

If the increase in selective advantage is defined as B_{xy} where

$$B_{xy} = b_{xy,\infty}/b_{xy,2}, \tag{12}$$

then by examination of equations (10) and (11) it can be seen that B_{xy} is independent of the sex of the ancestor (x) when $\tau_m = \tau_f$ as is the case for mass selection or selection using index I_D, whilst for selection using I_H, B_{xy} is dependent on x.

Evaluation of equation (8) with $t = 2$ rather than $t = \infty$. Let us now examine equation (8) but using, in the first instance, $t = 2$ rather than $t = \infty$. Firstly, by noting that $n_x = T/X = Fn_f/X$, it is found that

$$\begin{aligned}\mu^2_{r(xy),2} &= \frac{1}{4}\left\{\left(\frac{Y}{X}\right)^2 + b^2_{xy,2} V[A_{i(x)}]\right\} \\ &\approx \frac{1}{4}\left[\left(\frac{Y}{X}\right)^2 + \frac{1}{4\sigma^2_{I,2}} n_x^2 \tau_x^2 z_y^2 4\sigma^2_{Ax,1}\right] \\ &= \frac{1}{4}\left[\left(\frac{Y}{X}\right)^2 + \left(\frac{F}{X}\right)^2 n_f^2 \rho_{x,2} z_y^2\right]\end{aligned} \tag{13}$$

where $\rho_{x,2}$ is the correlation between full-sibs due to the breeding value of the parent of sex x, as defined in equation (5).

An extended form of equation (22) of Wray and Thompson (1990) approximates $\sigma^2_{r(xy),2}$ as,

$$\begin{aligned}\sigma^2_{r(xy),2} \approx \frac{1}{4} E\Bigg\{ n_x \left[p_y + \frac{b_{xy,2}}{n_x} A_{j(x)}\right]\left[1 - p_y - \frac{b_{xy,2}}{n_x} A_{j(x)}\right] \\ + f s_{xy,2} + h s_{xy,2}\Bigg\}\end{aligned} \tag{14}$$

where the first term is the binomial sampling variance of the number selected from the n_x offspring of sex y born to parents of sex x, each of which is selected with probability $p_y + (b_{xy,2}/n_x)A_{j(x)}$ which depends on the genetic merit of the parent. For the population structure considered here, this sampling should be hypergeometric because family sizes before selection are constant and the sampling is without replacement. Woolliams et al. (1993) approximated this by multiplying the term by $(X-1)/X$. For random selection ($b_{xy,2} = 0$), this results in $n_x p_y (1-p_y)(X-1)X^{-1} = YX^{-1}(T-Y)T^{-1}(X-1)X^{-1}$, this is an approximation to the exact hypergeometric variance $YX^{-1}(T-Y)(T-1)^{-1}(X-1)X^{-1}$, which shall be used here. (The correction to the terms involving $A_{j(x)}$ will be ignored until the section 'More accurate prediction of coselection of sibs'.) The term $f s_{xy,2}$ [equation (14)] is the probability of coselection of full-sibs not attributed to the parents of sex x [which has already been accounted for via the $b^2_{xy,2} V(A_{j(x)})$ in equation (13)], thus

$$f s_{xy,2} = \frac{F}{X} n_f (n_f - 1)(\rho_{D,2} - \rho_{x,2}) z_y^2$$

where $\rho_{D,2}$ is the correlation between indices of full-sibs [equation (3)] and $\rho_{x,2}$ is the correlation between indices of full-sibs attributed to the breeding value of the parents of sex x [equation (5)]. For a general correlation between sib indices ρ, ρz_y^2 is an approximation to the additional probability of coselection of two sibs of sex y (Robertson 1961). The coefficient $(F/X)n_f$ $(n_f - 1)$ reflects that each of the n_f offspring of sex y has $(n_f - 1)$ opportunities for coselection with a sib and that a parent of sex x contributes of F/X full-sib families. The term $h s_{xy,2}$ [in equation (14)] is the additional probability of coselection of half-sibs not attributed to the parent of sex x; $h s_{fy,2} = 0$ since in the hierarchical population structure there are no maternal half-sibs; and analogoulsy to the full-sib co-selection, if $\rho_{H,t}$ is the correlation between indices of half-sibs [equation (4)], then,

$$h s_{xy,2} = = \frac{F}{X}\left(\frac{F}{X} - 1\right) n_f^2 (\rho_{H,2} - \rho_{m,2}) z_y^2$$

since there are $(F/M)n_f$ offspring of a sire which have probabilities of coselection with their $((F/M) - 1)n_f$ half-sibs. This term was not introduced by Wray and Thompson (1990) or Woolliams et al. (1993) because under mass selection the covariance between indices of half-sibs is completely accounted for by the sire, i.e., $\rho_{H,2} = \rho_{m,2}$. Whilst this is true also under selection on an index of (individual and) full-sib records, it is not true for selection using an index of (individual full- and) half-sib records.

Following from the above we can write,

$$\mu^2_{r(xy),2} + \sigma^2_{r(xy),2} \approx \frac{1}{4}\left[\left(\frac{Y}{X}\right)^2 + \sigma^2_{e(xy)} + \sigma^2_{g(xy)}\right] \tag{15}$$

where,

$$\begin{aligned}\sigma^2_{e(xy)} &\approx \frac{Y}{X}\frac{T-Y}{T-1}\frac{X-1}{X} + \frac{F}{X} n_f (n_f - 1)(\rho_{D,2} - \rho_{x,2}) z_y^2 \\ &\quad + \frac{F}{X}\left(\frac{F}{X} - 1\right) n_f^2 (\rho_{H,2} - \rho_{x,2}) z_y^2,\end{aligned}$$

and (16)

$$\begin{aligned}\sigma^2_{g(xy)} &= \left[\frac{F}{X} n_f (n_f - 1) + \frac{F}{X}\left(\frac{F}{X} - 1\right) n_f^2\right] \rho_{x,2} z_y^2 \\ &= \left[1 - \frac{X}{F n_f}\right] b^2_{xy,2} V[A_{j(x)}].\end{aligned}$$

The first term of equation (15) is the mean squared under random selection. The variance $\sigma^2_{e(xy)}$ is hypergeometric sampling variance appropriate under random selection plus covariances due to coselection of sibs which are attributed to correlations arising from the mate of the parent of sex x or to shared estimation errors of family means. The variance $\sigma^2_{(g)xy}$ represents covariances of selection between sibs which are attributable to the parent of sex x. The form of $\sigma^2_{g(xy)}$ in terms of $b_{xy,2}$ will be used later.

The covariances of long-term contributions between male and female descendants [in equation (8) with $t = 2$ instead of ∞] can be written similarly except that there is no hypergeometric sampling term. By noting that each offspring has n_f full-sibs of the opposite sex, it is found that,

$$\mu_{r(xm),2}\mu_{r(xf),2} + \sigma_{r(xm,xf),2} \approx \frac{1}{4}\left[\frac{M}{X}\frac{F}{X} + \sigma_{e(xm,xf)} + \sigma_{g(xm,xf)}\right] \tag{17}$$

where

$$\sigma_{e(xm,xf)} = \frac{F}{X} n_f^2 (\rho_{D,2} - \rho_{x,2}) z_m z_f + \frac{F}{X}\left(\frac{F}{X} - 1\right) n_f^2 (\rho_{H,2} - \rho_{x,2}) z_m z_f$$

and (18)

$$\sigma_{g(xm,xf)} = \left(\frac{F}{X}\right)^2 n_f^2 \rho_{x,2} z_m z_f = b_{xf,2} b_{xm,2} V(A_{j(x)}).$$

Substituting equation (15) and (17) into (8) [ignoring for the moment that in (8) $t = \infty$] results in

$$\frac{1}{64M}\left\{4 + [\sigma^2_{e(mm)} + \sigma^2_{g(mm)}] + 2\left[\frac{M}{F}\right][\sigma_{e(mm,mf)} + \sigma_{g(mm,mf)}] + \left[\frac{M}{F}\right]^2 [\sigma^2_{e(mf)} + \sigma^2_{g(mf)}]\right\} + \frac{1}{64F}\left\{4 + [\sigma^2_{e(ff)} + \sigma^2_{g(ff)}] + 2\left[\frac{F}{M}\right][\sigma_{e(fm,ff)} + \sigma_{g(fm,ff)}] + \left[\frac{F}{M}\right]^2 [\sigma^2_{e(fm)} + \sigma^2_{g(fm)}]\right\}. \quad (19)$$

Evaluation of equation (8) with $t = \infty$. Finally let us consider the case when $t = \infty$. There are three important aspects to take into account:

(1) Under random selection (i.e., $b_{xy,t} = 0$) the mean of the long-term contributions of the ancestors is the same for $t = 2$ as for $t = \infty$ [see equation (9)] and so the first terms of equations (15) and (17) remain unchanged from $t = 2$ to $t = \infty$. The method of Woolliams et al. (1993) shows that this holds also under selection. [This is term A3.1 of Woolliams et al. (1993)]

(2) Under random selection, the collective contributions of the variances of the long-term relationships to rate of inbreeding is increased by a factor of 2 from $t = 2$ to $t = \infty$. This does not imply that each $\sigma^2_{r(xy)}$ doubles, but rather via dispersion of genes the total contribution of $\lambda_{x,t}$ for $t = \infty$ is twice that of $t = 2$, where

$$\lambda_{x,t} = \frac{1}{16X}\left[\left(\frac{X}{M}\right)^2 \sigma^2_{r(xm),t} + 2\frac{X}{M}\frac{X}{F}\sigma_{r(xm,xf),t} + \left(\frac{X}{F}\right)^2 \sigma^2_{r(xf),t}\right] \quad (20)$$

which are the variance and covariance terms in equation (8). This doubling has been observed in simulation when selection is at random. It is also as intrinsic to the predictions of Wray and Thompson (1990) and Woolliams et al. (1993) (in the method of the latter it arises from the summation of terms that occur each generation from $t = 2$ onwards, but weighted by 2^{t-2}). In selected populations, simulation results show that $\lambda_{x,\infty} > 2\lambda_{x,2}$ where $\lambda_{x,2}$ takes the form,

$$\lambda_{x,2} = \frac{1}{64X}\left\{\left(\frac{X}{M}\right)^2 [\sigma^2_{g(xm)} + \sigma^2_{e(xm)}] + 2\frac{X}{M}\frac{X}{F}[\sigma_{g(xm,xf)} + \sigma_{e(xm,xf)}] + \left(\frac{X}{F}\right)^2 [\sigma^2_{g(xf)} + \sigma^2_{e(xf)}]\right\}. \quad (21)$$

Woolliams et al. (1993) show that prediction of $\lambda_{x,\infty}$ can be achieved by $2\lambda_{x,2}$ plus extra terms considered in point 3 below. Under random selection ($b_{xy,2} = 0$) equations (21) and (20) are identical but, under selection, equation (21) also contains the $b^2_{xy,2}$ terms from equation (15). From Appendix 3 of Woolliams et al. (1993) terms A3.2, A3.6, A3.13, A3.16 and A3.17 sum to $2(\lambda_{m,2} + \lambda_{f,2})$ for mass selection.

(3) The doubling of (21) is insufficient to account for the cumulative selective advantage and extra terms must be included. The selective advantage from parent to offspring is included in the $\sigma^2_{g(xy)}$ and $\sigma_{g(xm,xf)}$ terms which are functions of $b_{xy,2}$ [equations (16) and (18)]. Investigation of the method of Woolliams et al. (1993) suggests that (by making some assumptions discussed below) the increase in selective advantage from ancestor to descendant can be accounted for by replacing $b_{xy,2}$ in $\sigma^2_{g(xy)}$ and $\sigma_{g(xm,xf)}$ by $b_{xy,\infty}$. Equivalently, this can be achieved by multiplying $\sigma^2_{g(xy)}$ by B^2_{xy} and $\sigma_{g(xm,xf)}$ by $B_{xm}B_{xf}$ where B_{xy} was defined in equation (12).

Accounting for these points results in the prediction of rate of inbreeding,

$$\Delta F \approx \frac{1}{32M}\left\{2 + \sigma^2_{e(mm)} + B^2_{mm}\sigma^2_{g(mm)} + 2\left(\frac{M}{F}\right)\cdot[\sigma_{e(mm,mf)} + B_{mm}B_{mf}\sigma_{g(mm,mf)}] + \left(\frac{M}{F}\right)^2 [\sigma^2_{e(mf)} + B^2_{mf}\sigma^2_{g(mf)}]\right\} + \frac{1}{32F}\left\{2 + \sigma^2_{e(ff)} + B^2_{ff}\sigma^2_{g(ff)} + 2\left(\frac{F}{M}\right)\cdot[\sigma_{e(fm,ff)} + B_{fm}B_{ff}\sigma_{g(fm,ff)}] + \left(\frac{F}{M}\right)^2 [\sigma^2_{e(fm)} + B^2_{fm}\sigma^2_{g(fm)}]\right\}.$$

Woolliams et al. (1993) also showed that the rate of inbreeding predicted from long-term contributions should be corrected for contributions from the base population. This correction also applies to equation (22) resulting in a final prediction ΔF where

$$\Delta F = \Delta F(1 + 2\Delta F). \quad (23)$$

Relationship to equation (4) of Woolliams et al. (1993)

Equivalent terms to those in Appendix 3 of Woolliams et al. (1993) have been derived for index selection (Wray and Woolliams, unpublished notes), but their form is complex. The complexity can be traced to the inequality between Cov(breeding value of sire, index or offspring) and Cov(breeding value of dam, index or offspring) i.e., $\tau_m \neq \tau_f$ for I_H. Under mass selection (and I_D) this asymetry between sexes does not exist. When terms involving τ_m and τ_f are multiplied and accumulated over generations many more types of terms result than in the analogous derivation for mass selection. This is illustrated by derivation of $b_{xy,\infty}$ for index selection in Appendix 2. Approximations invoked for index selection in point 3 above, involve using $\tau_m = \tau_f$ for some product terms. If the equations for $\sigma^2_{g(xy)}$, $\sigma_{g(xm,xf)}$, $\sigma^2_{e(xy)}$, $\sigma_{e(xm,xf)}$ and B_{xy} [equations (16), but approximating $YX^{-1}(T-Y)\cdot(T-1)^{-1}(X-1)X^{-1}$ to $YX^{-1}(T-Y)T^{-1}$, (18) and (12) are substituted into equation (22) then the following equality results

$$\Delta F \approx \frac{1}{8M}\left\{1 + i^2\left[\frac{M}{F}(\rho_{D,2} - \rho_{m,2}) + \left(1 - \frac{M}{F}\right)\cdot(\rho_{H,2} - \rho_{m,2}) + \rho_{m,2}Q^2_m\right]\right\} + \frac{1}{8F}[1 + i^2((\rho_{D,2} - \rho_{f,2}) + \rho_{f,2}Q^2_f)] - \frac{1}{32T}[4 + (i^2_m + i^2_f)((\rho_{D,2} - \rho_{f,2}) + (\rho_{D,2} - \rho_{m,2}) + 2i^2(\rho_{m,2}Q^2_m + \rho_{f,2}Q^2_f))], \quad (24)$$

where $Q_x = B_{xy}i_y/i$. Under mass selection (where $\rho_{D,2} = \rho_{m,2} + \rho_{f,2}$, $\rho_{H,2} = \rho_{m,2}$, $Q_x = S_\infty$), equation (24) reduces to equation (4)

of Woolliams et al. (1993) with their $K = i^2(S_\infty^2 - 1) + 2i[\frac{1}{2}(i_m c_m + i_f c_f) + i^2 S_\infty](1 - c^2)^{-1}$ approximated to $2i^2(S_\infty^2 - 1)$ and where their equation (4) has ignored terms in T except for $4/32T$.

Relationship to the equation of Latter (1959) and Hill (1979)

The form of equation (22) has been chosen for its similarity (and equality when B_{xy} is set to unity) to the discrete generation equation for the prediction of rate of inbreeding of Latter (1959) and Hill (1979) which is based on variance of family size in one generation for random selection. The variance of family size of selected offspring of sex y from parents of sex x, σ_{xy}^2, is equal to $\sigma_{e(xy)}^2 + \sigma_{g(xy)}^2$ here (and similarly for covariances). They derived their equation from a genetic drift argument, where effective population size is defined by the variance in change of gene frequency. Their expression was derived to account for non-genetic differences in fecundity and viability of offspring, $\sigma_{g(xy)}^2 = 0$, rather than for selection on a heritable trait, although it has been used as such (e.g., de Vries et al. 1990; Wray et al. 1990). The two-generation Latter-Hill equation proposed by Wray et al. (1990) is expected to be approximately equal to equation (22) but with $B_{xy} = b_{xy,3}/b_{xy,2}$. The Latter-Hill equation ignores some higher-order terms which may be approximately incorporated through the correction of equation (23). Like the equation (4) of Woolliams et al. (1993), equation (22) could be rewritten in the form of the (one-generation) Latter-Hill prediction of inbreeding plus a term describing the proliferation of lines from superior ancestors at the expense of their inferior contemporaries.

Relationship with the equation of Robertson (1961)

Equation (22) can also be related to the prediction of Robertson (1961) for populations of full-sib families ($M = F$). Understanding of his method has been hindered by an anomaly in the derivation whereby the interpretation of the N used changes from N = number of full-sib families (therefore the number of parents is $2N$) to N is the number of parents (Felsenstein 1989). However, the (one-generation) result can be derived using the method presented in Latter (1959). Robertson's prediction for one-generation can be obtained by setting $M = F = N$ in the Latter-Hill equation in which Poisson distribution of family size and sampling with replacement are assumed. (Also there is assumed to be no environmental correlation between full-sibs so that $\rho = \rho_D = \rho_m + \rho_f$). Robertson argued for a two-fold increase in selective advantage from generation 2 to infinity ($B = 2$), but which Wray and Thompson (1990) argued should be $B = B_{xy}$ in the notation of this paper. Robertson's prediction is,

$$\Delta F = \frac{1}{4N}(1 + B^2 i^2 \rho),$$

which based on a more thorough theoretical derivation equation (22) with $M = F$ reduces to,

$$\Delta F = \frac{1}{4N}\left[1 + \frac{1}{2}(1 + B^2) i^2 \rho\right].$$

More accurate prediction of coselection sibs

The use of $z_x z_y \rho$ [e.g., in $f s_{xy,2}$ and $h s_{xy,2}$ in equation (14)] is a first-order approximation to the probability of coselection of a pair of sibs of sexes x and y over and above that due to chance alone. This can be more-accurately predicted using the approximation of Mendell and Elston (1974),

$$\left\{\Phi\left[\frac{i_x \rho - v_y}{(1 - k_x)^{1/2}}\right] p_x - p_x p_y\right\}.$$

This expression is well-defined for $x = y$. However, when $x \neq y$, whilst the probability is symmetric in x and y, the expression is not and both forms are approximations to it. Mendell and Elston (1974) show that accuracy decreases with i so the preferred form has $x = m$ and $y = f$. Under index selection when correlations between sib indices can become very high, the use of this more-accurate prediction of coselection of sibs is important.

When $x = y$, both first-order and second-order approximations to the probability should be multiplied by $((Y-1)/Y)(T/(T-1))$ in an attempt to account for selection without replacement.

Populations in which family size is large relative to the number selected

In the predictions of variance of long-term contributions (or variance of one-generation family size) discussed above, variances have been increased over and above random selection by considering coselection of sibs. This has been calculated as a probability of selection of a pair of sibs multiplied by the possible number of pairs available for selection, without imposition of a constraint of total number selected. In general, this approximation is good (see Results section, Table 2), but when family size available for selection is greater than the number selected, for example $(F/M)n_f > M$, then highly-inflated probabilities of coselection and variances can arise, particularly when the correlation between selection criterion of sibs is high. At the extreme, if $\rho_H = 1$, then all M males will be selected from a single half-sib family. Wray et al. (1990) discussed this problem and for these situations proposed the use of p_m' instead of p_m, where $p_m' = (1 - \rho_H)p_m + \rho_H(F/M)n_f/T$. The full impact of this approximation affects several of the equations presented in this paper and their adapted form is given in Appendix 3.

Simulation

Predictions from equation (23) are compared to rates of inbreeding observed from simulation. Simulations for mass selection are those presented in Wray and Thompson (1990) based on 100 replicates. For index selection, simulations are similar except that selection is based on either I_H or I_D. Populations have $M = 20$ males, $F = 20, 40, 200$ females with $n_f = 3, 6$ offspring of each sex per dam. Heritabilities considered are $h^2 = 10^{-6}, 0.1, 0.2, 0.4, 0.6, 0.99$, $\sigma_{P,0}^2 = 1$ and common environment variance $\sigma_c^2/\sigma_P^2 = 0$. Heritability values close to zero and close to unity have been investigated so that the predictions can be tested at the extremes where it is possible to postulate the way in which selection is operating. When heritability is exactly zero, index weights are null and selection is at random. But when heritability is close to zero ($h^2 = 10^{-6}$) selection on I_H (or I_D) is close to selection on the family mean since the correlation between sib indices is high. When $h^2 = 0.99$, selection using any of the three indices will result in selection of the same individuals. Other populations simulated have $M = F = 20$, with $n_f = 3, 6, 12, 20$ under mass selection for traits with $h^2 = 10^{-6}$, $\sigma_{P,0}^2 = 1$ but with $\sigma_c^2 = 0.00, 0.20, 0.60, 1 - 10^{-6}$. Within the simulations many statistics are calculated which are checked with predictions. These include variances of breeding values of selected ancestors (about the mean of the selected group), correlations between selection criteria of sibs (born in generation 2, calculated by analysis of variance), probabilities of coselection of sibs, variance of family size from parents (born in generation 1) to offspring, $b_{xy,t}$, total sums of squares of long-term contributions and rates of inbreeding. Rates of inbreeding presented are the average of those observed from generations 5 to 14. Simulation results are the average of 1000 replicates for $F = 20$, 40 and 500 replicates for $F = 200$. An example calculation is given in Appendix 4.

Results

Variances and correlations

Predicted values and prediction errors of variances of true breeding values of selected individuals born in generation 1 about the mean of the selected group $[V(A_{j(x)})]$ are presented in Table 1. Predictions agree well with simulated values (maximum error 10%). In comparison, calculation of the variance of breeding values about the unconditional mean $[V(A^*_{j(x)})]$ can lead to overestimation by as much as 20% (data not shown). Correlations between full and half-sib index values ($\rho_{D,2}$ and $\rho_{H,2}$) are also presented in Table 1; predicted values are those described in the notation section using $\sigma^2_{Ax,1}$ and calculated by Appendix 2. Predictions of correlations are also accurate (maximum standard error of simulations is 0.007) although correlations of 0.8 or greater tend to be underpredicted. Correlations shown are between sibs born in generation 2; these correlations may be substantially lower (particularly for high h^2) than correlations between sibs born in generation 1 (before selection).

Predicted values and prediction errors of variances of family size (of offspring born in generation 2 from male parents) are presented in Table 2. Predictions of (co)variances of family size from female parents show smaller prediction errors (data not shown). Probabilities of coselection were also examined but these show a similar pattern to the variances of family size. There is a tendency to overpredict (co)variances of family size from male parents when selection uses I_H and in situations where n_f and ρ_H are high (i.e., h^2 low). The overprediction becomes particularly acute in situations where $(F/M)n_f > M$ and ρ_H is high. Predictions using p'_m and the equations of Appendix 3 are also presented, which do remarkably well given that the adjustments are based on heuristic arguments.

Expected long-term contributions

Predicted values and prediction errors of $b_{mm,2}$ and B_{xy} [equations (11) and (12)] are presented in Table 3. When selection uses I_D (or I_P, data not shown) the simulation results are in good agreement with the theoretical result that B_{xy} is independent of x, the sex of the ancestor (since $\tau_m = \tau_f$). This is not true for selection using I_H, where for hierarchical populations, breeding values of male ancestors are more highly correlated to their offspring's index values than are female ancestors ($\tau_m > \tau_f$). If this difference is ignored and an average τ is used when selection is on I_H then serious errors in the prediction of rate of inbreeding and its components arise (data not shown). Simulation results show that the increase in long-term contributions from generation 2 to ∞ is greater for female ancestors ($B_{fy} > B_{my}$), this is expected from evaluation of equation (12) which can be shown to be a function of τ/τ_x. For two populations with the same structure and heritability, but where selection has used different indices, it is found that in the population where $b_{xy,2}$ is higher, then B_{xy} is lower. The predictions of $b_{xy,2}$ are generally good (shown only for $b_{mm,2}$), although they tend to underpredict when $h^2 = 0.99$. For populations where $(F/M)n_f > M$, predictions of $b_{mm,2}$ are too high. Prediction of $b_{mm,2}$ using the results which depend on p'_m give satisfactory predictions. Predictions errors in B_{xy} are found to be robust compared to prediction errors in $b_{xy,2}$ and $b_{xy,\infty}$.

Rates of inbreeding

Predicted values and prediction errors for rates of inbreeding are presented in Table 4. In some simulations, particularly for high h^2, rates of inbreeding were observed to be somewhat higher in generation 2 (and sometimes 3). For example, when $h^2 = 0.99$, the first round of selection (where selected individuals tend to come from a few good families) results in a high initial rate of inbreeding. In subsequent generations, the rate of inbreeding is less as a result of selection of the best individuals across families that are genetically less variable. However, in all cases investigated, inbreeding reached an approximately steady rate by generation 4 and over the generations included in the average. For selection on I_H, the prediction error of the rate of inbreeding, as calculated from the total sum of squares of long-term contributions, accumulated within the simulation, [equation (6) with the correction for base contributions, equation (23)] is also presented. This demonstrates that under index selection where rates of inbreeding can be much higher than under mass selection, the prediction of the rate of inbreeding via long-term contributions remains appropriate.

Predictions of rate of inbreeding for mass selection are accurate when $F \geq 100$ (maximum error of prediction 7%), but have a tendency to underpredict for $F \leq 40$ (maximum error 6%). This is the same pattern as found by the approximation equation (4) of Woolliams et al. (1993). (Mass selection predictions presented here use the Mendell and Elston (1974) probabilities of coselection, whereas Woolliams et al. used only the first order approximation for their tabulated results.) Predictions are also accurate for selection on I_D but with a tendency to overpredict at low h^2 (high ρ_F) (which partially reflects overprediction of one-generation variance of family size, data not shown), with maximum errors of 6% for $n_f = 3$ and 13% for $n_f = 6$. The predictions for I_H are accurate providing that h^2 is greater than 0.2 and $M > (F/M)n_f$, but may overestimate otherwise. The overprediction can be as much as 114% ($M = 20$, $F = 200$, $n_f = 6$,

Table 1. Variance of breeding values of selected sires, $V[A_{j(m)}]$, and dams, $V[A_{j(f)}]$, and correlations between index values of full sibs, $\rho_{F,2}$ and half-sibs, $\rho_{H,2}$ (born generation 2). Predicted (P), predicted-simulated (P-S) and percentage error [100*(P-S)/S] values for populations with $M = 20$ males, F females and n_f offspring of each sex born/female with selection using indices I_D or I_H

I	F	n_f	h^2	$V[A_{j(f)}]$			$V[A_{j(m)}]$			$\rho_{F,2}$			$\rho_{H,2}$		
				P	P-S	% Error	P	P-S	% Error	P	P-S	% Error	P	P-S	% Error
I_D	40	3	0.00[a]							0.889	−0.038	−4.1	0.000	0.000	0.0
			0.10	0.083	0.000	0.0	0.078	+0.001	+1.3	0.852	+0.010	+1.2	0.086	−0.01	−10.4
			0.40	0.242	0.000	0.0	0.219	−0.002	−0.9	0.694	+0.009	+1.3	0.164	−0.005	−3.0
			0.99	0.273	0.000	0.0	0.195	0.000	0.0	0.197	+0.002	+1.0	0.080	0.000	0.0
I_H	40	3	0.00[a]							0.905	−0.002	−0.2	0.667	−0.002	−0.3
			0.10	0.081	0.000	0.0	0.076	0.000	0.0	0.861	+0.002	+0.2	0.573	+0.007	+1.2
			0.40	0.240	+0.013	+5.7	0.217	+0.002	+0.9	0.692	+0.008	+1.2	0.355	+0.002	+0.6
			0.99	0.273	0.000	0.0	0.195	+0.005	+2.6	0.197	+0.005	+2.6	0.081	+0.002	+2.5
		6	0.00[a]							0.944	−0.010	−1.1	0.722	−0.006	−0.8
			0.10	0.073	0.000	0.0	0.067	0.000	0.0	0.888	−0.001	−0.1	0.573	+0.006	+1.1
			0.40	0.211	+0.006	+2.9	0.194	+0.004	+2.1	0.679	+0.011	+1.6	0.330	−0.012	−3.5
			0.99	0.199	0.000	0.0	0.155	+0.006	+4.0	0.156	+0.003	+2.0	0.068	+0.001	+1.5
	200	3	0.00[a]							0.956	−0.032	−3.2	0.844	−0.023	−2.7
			0.10	0.080	+0.001	+1.3	0.067	+0.002	+3.0	0.882	−0.020	−2.2	0.637	−0.010	−1.5
			0.40	0.242	−0.002	−0.8	0.198	+0.004	+2.1	0.681	−0.009	−1.3	0.330	−0.006	−1.8
			0.99	0.280	+0.003	+1.1	0.128	+0.013	+10.1	0.176	+0.008	+4.8	0.054	+0.003	+5.9
		6	0.00[a]							0.976	−0.023	−2.3	0.881	−0.029	−3.1
			0.10	0.073	+0.001	+1.4	0.060	−0.002	−3.2	0.893	−0.021	−2.1	0.591	−0.015	−2.5
			0.40	0.216	+0.003	+1.4	0.182	+0.001	+0.5	0.671	−0.006	−0.9	0.308	−0.003	−1.0
			0.99	0.205	0.000	0.0	0.113	+0.009	+8.0	0.143	0.003	+2.1	0.050	+0.003	+6.4

[a] $h^2 = 10^{-6}$

Table 2. Variances and covariances of family sizes of selected offspring from male parents. Predicted (P), predicted-simulated (P-S) and percentage error [100*(P-S)/S] values for populations with $M = 20$ males, F females and n_f offspring of each sex per female with selection using indices I_D or I_H

I	F	n_f	h^2	σ^2_{mm}[a]				σ^2_{mf}[a]			$\sigma_{mm,mf}$[a]			
				P	P-S	% Error	()[b]	P	P-S	% Error	P	P-S	% Error	()[b]
I_D	40	3	0.00[c]	1.87	+0.02	+1.1		3.09	+0.05	+1.6	1.79	+0.05	+2.9	
			0.10	1.88	+0.03	+1.6		3.18	+0.06	+1.9	1.86	+0.07	+3.9	
			0.40	1.70	+0.04	+2.4		2.94	+0.10	+3.5	1.62	+0.09	+5.9	
			0.99	1.04	−0.07	−6.3		1.78	−0.04	−2.2	0.47	−0.06	−11.1	
I_H	40	3	0.00[c]	2.91	+0.07	+2.5		4.96	+0.08	+1.6	3.11	+0.09	+3.0	
			0.10	2.62	+0.03	+1.2		4.51	−0.10	+2.3	2.81	+0.10	+3.7	
			0.40	1.96	+0.02	+1.0		3.43	0.01	+0.3	1.97	+0.07	+3.7	
			0.99	1.04	−0.09	−8.0		1.78	−0.05	−2.7	0.47	−0.07	−13.0	
		6	0.00[c]	6.32	+0.57	+9.9		12.39	+0.63	+5.4	7.73	+0.55	+7.7	
			0.10	5.18	+0.46	+9.7		10.38	+0.53	+5.4	6.48	+0.53	+8.9	
			0.40	3.25	+0.18	+5.9		6.79	+0.18	+2.7	3.96	+0.29	+7.9	
			0.99	1.23	−0.12	−8.9		2.51	−0.21	−7.8	0.67	−0.18	−21.1	
	200	3	0.00[c]	15.00	+4.62	+44.5	(+6.3)	131.34	+5.43	+4.3	19.74	+0.60	+3.1	(−2.1)
			0.10	8.89	+1.84	+26.1	(+3.8)	93.57	+2.61	+2.9	17.03	+0.91	+5.6	(−0.6)
			0.40	3.83	+0.38	+11.0	(+3.8)	49.67	−0.09	−0.2	9.65	+0.47	+5.1	(+1.2)
			0.99	1.25	−0.14	−10.0	(−10.8)	13.67	−0.62	−4.3	1.67	−0.34	−16.9	(−17.4)
		6	0.00[c]	31.56	+18.83	+147.9	(+1.9)	337.30	+23.03	+7.3	49.22	+4.86	+11.0	(−13.3)
			0.10	13.67	+5.49	+67.1	(+2.9)	194.87	+9.43	+5.1	35.20	+5.20	+17.3	(−7.5)
			0.40	5.28	+1.18	28.8	(+7.1)	97.02	+2.83	+3.0	17.41	+2.16	+14.2	(+0.8)
			0.99	1.33	−0.14	−9.5	(−10.9)	21.10	−1.29	−5.8	2.38	+0.34	+16.7	(+14.2)

[a] σ^2_{xy} and $\sigma_{xm,xf}$ are predicted as $\sigma^2_{g(xy)} + \sigma^2_{e(xy)}$ and $\sigma_{g(xm,xf)} + \sigma_{e(xm,xf)}$ respectively

[b] Percentage errors in parenthesis are achieved when using p'_m instead of p_m, see section 'Population in which family size is large relative to number selected'

[c] $h^2 = 10^{-6}$

Table 3. One-generation selective advantage of male ancestors to their male offspring born in generation 2, $b_{mn,2}$ [equation (11)] and increase in selective advantage from one-generation to long-term for ancestors of each sex to descendants of each sex, B_{xy} [equation (12)]. Predicted (P), predicted-simulated (P-S) and percentage error [100* (P-S)/S] values for populations with $M = 20$ males, F females and n_f offspring of each sex/female and with selection using indices I_D or I_H

I	F	n_f	h^2	$b_{mm,2}$				B_{mm}			B_{fm}			B_{mf}			B_{ff}		
				P	P-S	%Error	()[a]	P	P-S	%Error	P	P-S	%Error	P	P-S	%Error	P	P-S	%Error
I_D	40	3	0.10	1.58	+0.03	+1.9		1.42	−0.06	−4.4	1.43	+0.03	+2.1	1.96	+0.09	+ 4.8	1.96	+0.03	+ 1.6
			0.40	1.30	−0.02	−1.5		1.13	−0.02	−1.7	1.13	−0.04	−3.4	1.55	−0.04	− 2.5	1.55	−0.02	− 1.3
			0.99	0.96	−0.08	−7.6		0.98	−0.01	−1.0	0.98	−0.02	−2.0	1.35	−0.07	− 4.9	1.35	−0.15	−10.0
I_H	40	3	0.10	1.94	0.00	0.0		1.31	+0.08	+6.5	1.51	+0.06	+4.1	1.80	+0.12	+7.1	2.07	+0.15	+7.8
			0.40	1.43	−0.02	−1.3		1.08	−0.02	−1.8	1.16	−0.05	−4.1	1.49	−0.04	− 2.6	1.59	−0.06	− 3.6
			0.99	1.03	−0.07	−6.8		0.98	−0.04	−3.9	0.98	−0.06	−5.7	1.35	−0.11	− 7.5	1.35	−0.14	− 9.3
		6	0.10	3.08	+0.05	+1.6		1.25	+0.05	+4.2	1.42	+0.07	+5.2	1.53	+0.09	+6.3	1.74	+0.09	+5.5
			0.40	2.01	+0.04	+2.0		1.06	−0.02	−1.9	1.11	−0.04	−3.4	1.30	+0.01	+0.8	1.37	−0.04	− 2.8
			0.99	1.21	−0.16	−11.7		1.00	+0.02	+2.0	1.00	0.00	0.0	1.22	0.00	0.0	1.23	−0.07	− 5.4
	200	3	0.10	4.87	+0.45	+10.2	(+4.8)	0.92	+0.01	+1.1	1.60	+0.07	+4.2	1.87	+0.19	+11.3	3.27	+0.14	+4.5
			0.40	2.58	+0.06	+2.4	(−0.3)	0.84	−0.01	+1.1	1.04	−0.06	−5.4	1.72	0.00	0.0	2.12	−0.16	− 7.0
			0.99	1.45	−0.30	−17.1	(−17.1)	0.83	+0.01	+1.2	0.83	0.00	0.0	1.69	−0.30	−15.1	1.70	−0.39	−18.7
		6	0.10	6.25	+0.88	+16.4	(+2.6)	0.93	−0.03	−3.1	1.46	+0.12	+9.0	1.54	+0.26	+18.8	2.42	+0.05	+2.1
			0.40	3.05	+0.20	+7.0	(−0.7)	0.88	−0.03	−3.2	1.01	−0.01	−1.0	1.46	+0.07	+5.0	1.68	−0.06	− 3.4
			0.99	1.65	−0.27	−14.1	(−15.1)	0.87	+0.01	+1.2	0.87	0.00	0.0	1.44	−0.10	− 6.5	1.45	−0.27	−15.7

[a] Percentage errors in parenthesis are achieved when using p'_m instead of p_m, see section 'Populations in which family size is large relative to number selected'

Table 4. Rate of inbreeding × 100. Predicted (P), predicted-simulated (P-S) and percentage errors [100*(P-S)/S] values for populations with $M = 20$ males and F females and n_f offspring of each sex/female and with selection using I_P, I_D or I_H. For selection using I_H, (S*-S) represents the difference between rate of inbreeding calculated from equation (23) using ΔF from equation (6) (where $\sum r^2$ is calculated in the simulation) and rate of inbreeding from inbreeding coefficients (calculated in the simulation)

F	n_f	h^2	I_P			I_D			I_H				
			P	P-S	%Error	P	P-S	%Error	P	S*-S	P-S	%Error	()[a]
20	3	0.00[b]	1.05	−0.02	−1.9	2.57	+0.14	+5.8	2.57	−0.03	+0.14	+5.8	
		0.10	1.18	−0.05	−4.0	2.68	+0.07	+2.7	2.68	+0.01	+0.07	+2.7	
		0.20	1.26	−0.07	−5.2	2.58	+0.03	+1.2	2.58	−0.01	+0.03	+1.2	
		0.40	1.36	−0.06	−4.2	2.29	0.00	0.0	2.29	+0.03	0.00	0.0	
		0.60	1.41	−0.09	−6.0	1.99	−0.02	−1.0	1.99	−0.01	−0.02	−1.0	
		0.99	1.37	0.00	0.0	1.38	+0.01	+0.7	1.38	+0.04	+0.01	+0.7	
40	3	0.00[b]	0.84	+0.01	+1.2	1.92	+0.07	+3.8	2.42	−0.04	+0.14	+6.1	
		0.10	0.97	−0.01	−1.0	2.14	+0.03	+1.4	2.52	0.00	+0.11	+4.6	
		0.40	1.14	−0.04	−3.3	1.89	−0.03	−1.6	2.01	0.00	+0.02	+1.0	
		0.99	1.09	+0.01	+0.9	1.10	+0.02	+1.9	1.10	+0.07	+0.02	+1.9	
	6	0.00[b]	0.89	+0.01	+1.1	4.02	+0.46	+12.9	5.28	−0.08	+0.82	+18.4	
		0.10	1.11	−0.06	−5.1	4.37	+0.36	+9.0	5.11	−0.08	+0.62	+13.8	
		0.40	1.42	−0.08	−5.3	3.20	+0.07	+2.2	3.36	+0.04	+0.18	+5.7	
		0.99	1.26	+0.07	+7.4	1.27	+0.05	+4.1	1.27	+0.08	+0.04	+3.3	
200	3	0.00[b]	0.67	+0.01	+1.5	1.16	+0.03	+2.7	4.27	−0.03	+1.34	+45.7	(+20.5)
		0.10	0.81	−0.01	−1.2	1.56	+0.03	+2.0	3.31	−0.01	+0.74	+28.8	(+16.3)
		0.20	0.90	+0.04	+4.7	1.60	−0.01	−0.6	2.58	+0.03	+0.34	+15.2	(+8.0)
		0.40	0.99	−0.01	−1.0	1.48	−0.03	−2.0	1.84	+0.02	+0.08	+4.5	(+1.1)
		0.60	0.99	0.00	0.0	1.29	−0.01	−0.8	1.41	+0.01	+0.01	+0.7	(−0.7)
		0.99	0.83	+0.02	+2.5	0.84	+0.02	+2.4	0.84	+0.30	0.00	0.0	(0.0)
	6	0.00[b]	0.68	0.00	0.0	2.11	+0.18	+9.3	9.65	−0.07	+5.13	+113.5	(+22.6)
		0.10	0.89	0.02	+2.2	2.81	+0.18	+6.8	5.58	+0.08	+1.90	+51.6	(+16.8)
		0.20	1.03	0.00	0.0	2.71	+0.10	+3.8	4.04	−0.06	+0.84	+26.3	(+6.9)
		0.40	1.15	−0.02	−1.7	2.26	+0.07	+3.2	2.66	+0.01	+0.30	+12.7	(+3.4)
		0.60	1.14	−0.05	−4.2	1.77	+0.02	+1.2	1.89	−0.05	+0.09	+5.0	(0.0)
		0.99	0.89	+0.02	+2.3	0.90	+0.02	+2.3	0.90	+0.06	+0.02	+2.3	(+2.3)

[a] Percentage errors in parentheses are achieved when using p'_m instead of p_m, see section 'Populations in which family size is large relative to number selected'

[b] $h^2 = 10^{-6}$

$h^2 = 10^{-6}$). The use of p'_m leads to improved predictions with maximum errors of 21% for $n_f = 3$ and 23% for $n_f = 6$.

Errors in the prediction of the rate of inbreeding

Errors in the prediction of the rate of inbreeding for index selection are greatest when h^2 is close to zero. In this case equation (22) reduces to the Latter-Hill equation. As such it is independent of any errors in prediction of selective advantage and depends only on the variance of one-generation family size. Prediction of the variance of family size is fairly good when the equations of Appendix 3 are used for populations of large family size relative to the number selected. Indeed, if variances of family size from the simulation are substituted into the Latter-Hill equation then overprediction of observed rate of inbreeding is found; for example for $I_H, F = 200$, $n_f = 6$ the predicted rate of inbreeding using variances from the simulation is 0.0558, which is close to the prediction using predicted variances of 0.0560 and which both overpredict the observed rate of inbreeding of 0.0452. (Even without the correction used here, equation (23), the Latter-Hill equation still overpredicts at 0.0508.) When $M = F$, index selection for a trait with near zero heritability, is equivalent to mass selection for a trait with near zero heritability and non-zero σ_C^2/σ_p^2. Further investigation of the problem was conducted with simulations of populations of this type; $\sigma^2_{(r)xy,t}$ and $\sigma_{(r)xm,xf,t}$ from these are presented in Table 5. According to the theory used in this paper $\lambda_{x,\infty}/\lambda_{x,2} = 2$ where $\lambda_{x,t}$ is defined in equation (20). However, it is found that when σ_C^2/σ_p^2 is high, the ratio is considerably less than 2. In this case selected parents are all chosen from a minimum of families. The problem is found to be more apparent when family size is large relative to the number selected. Indeed, in the most extreme case considered, when $h^2 = 10^{-6}$, $\sigma_C^2/\sigma_p^2 = 1-10^{-6}$ and $M = F = n_f = 20$, all offspring are chosen from a single family and the variance of contributions from ancestors to descendants cannot increase over and above that from parents to offspring. The ratio in Table 5 is shown to be unity. In this example, the observed rate of inbreeding is 0.191

Table 5. Variances of long-term contributions (from simulation) when $h^2 = 10^{-6}$ $\sigma_P^2 = 1$ and σ_C^2 is non-zero from simulation, $M = F = 20$ and with selection using I_P

σ_C^2	$n_f = 3$				$n_f = 6$				$n_f = 12$				$n_f = 20$			
	$\sigma^2_{r(xy),2}$	$\sigma_{r(xm,xf),2}$	$\sigma^2_{r(xy),\infty}$ [a]	Ratio [b]	$\sigma^2_{r(xy),2}$	$\sigma_{r(xm,xf),2}$	$\sigma^2_{r(xy),\infty}$ [a]	Ratio [b]	$\sigma^2_{r(xy),2}$	$\sigma_{r(xm,xf),2}$	$\sigma^2_{r(xy),\infty}$ [a]	Ratio [b]	$\sigma^2_{r(xy),2}$	$\sigma_{r(xm,xf),2}$	$\sigma^2_{r(xy),\infty}$ [a]	Ratio [b]
0.00	0.16	0.00	0.15	1.99	0.20	0.00	0.19	2.00	0.22	0.00	0.21	2.01	0.23	0.00	0.22	2.01
0.20	0.20	0.06	0.25	2.00	0.29	0.11	0.38	2.00	0.39	0.19	0.53	1.96	0.44	0.24	0.64	2.05
0.60	0.29	0.19	0.46	1.97	0.54	0.42	0.87	1.88	0.89	0.75	1.39	1.83	1.18	1.02	1.78	1.86
1.00 [c]	0.48	0.48	0.86	1.81	1.15	1.15	1.84	1.60	2.35	2.35	3.15	1.34	4.74	4.74	4.75	1.00

[a] $\sigma_{r(xy),\infty} = \sigma_{r(xm,xf),\infty}$ when $M = F$.
[b] Ratio $= \lambda_{x,\infty}/\lambda_{x,2}$ where $\lambda_{x,t}$ is defined in equation [20]
[c] $\sigma_C^2 \doteq 1-10^{-6}$

(after an initial rate of 0.250), identical to the classical result for full-sib mating (Wright 1931). The rate of inbreeding calculated from the observed long-term contributions [equation (6) with correction (23)] is 0.161. Whilst the rate of inbreeding calculated from observed or predicted (assuming family selection) variances of family size using the equation of Hill (1979) [which is equivalent to the prediction using equation (22)] is 0.243 [or 0.361 after correction by equation (23)].

Discussion

The prediction of rate of inbreeding presented here can be divided into three steps:

(1) Equality of rate of inbreeding calculated by identity of descent to equation (6) plus (23).
(2) Equality of equation (22) to (6) which includes (i) the assumption (used directly in this paper, but also implicit in Wray and Thompson 1990 and Woolliams et al. 1993) that the contribution of $\lambda_{x,t}$ [equation (21)] to rate of inbreeding increases by a factor of 2 or more from $t = 2$ to $t = \infty$, and (ii) the approximation of terms which describe the contributions of the expected long-term selective advantage when selection is on a heritable trait.
(3) Prediction of the components of equation (22), i.e., variances of one-generation family size and the long-term selective advantage terms.

Errors in prediction of the rate of inbreeding defined by identity by descent can occur in any of the three steps. The extreme example of $M = F = n_f = 20$, $h^2 = 10^{-6}$, $c^2 = 1-10^{-6}$ with mass selection highlights the first of these errors. In this example, selection is for the best full-sib family and the rate of inbreeding is that appropriate to repeated full-sib mating. It shows that the errors in the prediction of rate of inbreeding are not confined to the situation of selection on a heritable trait but can be concerned with constraints of population structure. The rates of inbreeding defined by squared contributions and by identity by descent may not be identical in all circumstances since approximations are invoked in the proof of equivalence given by Wray and Thompson (1990). However, for the range of breeding programmes investigated in Table 4, which includes some population structures which could be considered extreme for livestock, the equality between identity by descent rate of inbreeding and equations (6) plus (23) is good and the prediction of long-term contributions as a means of predicting the rate of inbreeding remains an appropriate goal.

In some situations, errors in the prediction of rate of inbreeding occur at stage 2, where the ratio $\lambda_{x,\infty}/\lambda_{x,2}$ [equation (21)] is less than 2. In the extreme example

above, of selection of a single full-sib family, this ratio is 1, as the variance of long-term contributions from ancestors to descendants achieves its maximum after a single generation. Since it is at stage 2 that the equality of equation (22) and the Latter-Hill equation is found (when selection is for a non-heritable trait), then errors arising at stage 2 are equally applicable to the Latter-Hill equation. Such errors have already been discussed by Wray et al. (1990) in comparison of the Latter-Hill prediction of inbreeding with the approach using the maximum eigen value from transition matrices (Woolliams 1989). Errors in stage 2 are less easy to detect explicitly when selection is on a heritable trait, but are likely to contribute to the prediction errors observed in Table 4, particularly when the family size available for selection is large relative to the number of parents and when the correlation between selection criteria of sibs is high.

Errors also occur at stage 3, in the prediction of variances of one-generation family size (Table 2) and long-term selective advantage (Table 3) (errors in the latter reflect errors in the former). The prediction of variances of family size assumes selection across families, but in examples when family size prior to selection is large and correlation between sib records is high, selection tends to be for selection of the best families, in which case the variance of family size after selection is less than predicted. A method has been presented to account for selection of half-sib families, but selection of the best full-sib families within half-sib families has been ignored.

For the populations investigated in this paper (Table 4), rates of inbreeding are underpredicted when selection is on phenotype alone, this is expected as the terms presented here are an approximation to the derivation of Woolliams et al. (1993). Under mass selection, the errors in steps 1–3 discussed above, are unlikely to occur for population structures relevant to livestock breeding. However, these errors can become important when selection is on an index which includes records of collateral relatives, ensuring that the correlation between the selection criteria of sibs is high. In these cases, there is a tendency for overprediction of ΔF, which is found despite the fact that (small positive) terms are ignored in the derivation. The correction to the rate of inbreeding for base contributions (equation (23)] sometimes causes a good prediction of ΔF to become an overprediction, the good prediction before the correction can be attributed to compensatory errors.

In summary, whilst the goal of prediction of rate of inbreeding via long-term contributions remains valid, the method presented here to achieve this prediction does not fully acount for constraints upon the variance of long-term contributions arising from the population structure. Under most circumstances the absence of such a constraint is not an issue, but it becomes important when correlations between sib indices are high (> 0.8) and family size is large relative to the number selected.

In many practical breeding programmes, a restriction is placed on the number of offspring selected per full-sib and half-sib family. Such a restriction would influence the prediction of $\sigma^2_{r(xy),2}$, $\sigma^2_{g(xy)}$, $\sigma^2_{e(xy),2}$ and the equivalent covariances, and the linear prediction of number of offspring selected per parent. However, the problems in the assumption of the increase of long-term contributions from $t = 2$ to $t = \infty$ are unlikely to arise.

When selection in a hierarchical population is on an index which includes half-sib records, a sire is more highly correlated to the index values of his offspring than are his mates. This results in an asymmetry between the contributions from male and female ancestors to the evolution of the rate of inbreeding which does not arise under mass selection or when full-sib records alone are included in an index. It is important to account for this asymmetry in the prediction of inbreeding when selection is on I_H. One-generation selective advantage ($b_{xy,2}$) is greater for I_H vs I_D and I_D vs I_P, whereas the increase in selective advantage (B_{xy}) shows the reverse pattern. Thus, one-generation predictions of rates of inbreeding are expected to underpredict the observed rate of inbreeding to a lesser extent when selection uses sib records compared to mass selection (examples are given in Wray 1989). Under mass selection and with no common environmental effects, the correlation between the selection criteria of sibs is entirely of genetic orgin, whilst under selection on family indices, the correlation is partly of environmental origin. Thus, the relative contribution to the total inbreeding of the long-term increase of the genetic component must be smaller for selection using family indices (Wray et al. 1990).

Prediction of rates of inbreeding when selection is based on estimated breeding values calculated by BLUP (best linear unbiased prediction) could be achieved using a selection index, an extension of I_H but including estimated breeding values of the sire, dam and other mates of the sire (Wray and Hill 1989). However, overprediction of rates of inbreeding will be expected as BLUP induces even higher correlations between sib indices than I_H. Use of the predictions obtained here for I_H, although not exact, would be an improvement on what has been used in the past.

This paper represents the completion of the second stage towards the joint description of progress and inbreeding in terms of the same predictable parameters. Woolliams et al. (1993) derived terms which describe the expected proliferation of lines and show how inbreeding can be related concisely to these terms in addition to the variance of family size in mass

selection. This paper extends this theory to the important case of index selection and could be adapted to other situations, such as sex-limited traits, non-hierarchical population structures, and variances in physical family size prior to selection (the subject of a later paper). It is now possible for a wide range of circumstances to assess quanta of information for their value in promoting progress and their value in promoting inbreeding.

Appendix 1. Expression for ΔF partitioned for male amd female ancestors

Under the assumption of constant rate of inbreeding each generation, Wray and Thompson (1990) showed that the rate of inbreeding can be expressed as,

$$\Delta F \approx \frac{1}{4(M+F)^2}\mathbf{1}^T\mathbf{C}_{1(1.t)}^T\mathbf{C}_{1(1.t)}\mathbf{1} = \frac{1}{4(M+F)^2}\sum_{i=1}^{M+F} r_i^2$$

where $\mathbf{C}_{1(1.t)}$ is a square matrix of order $M \times F$ of relationships between ancestors born in generation 1 and their descendants born in generation t, which result from the Mendelian samplings received by the ancestors. The elements of $\mathbf{C}_{1(1.t)}^T\mathbf{C}_{1(1.t)}$ are all the same when t is large, thus the pre- and post-multiplication by the unity vector 1 sums all the elements and division by $(M+F)^2$ gives the average. If the order of ancestors is males then females and the order of descendants is males then females, then $\mathbf{C}_{1(1.t)}$ can be partitioned as,

$$\mathbf{C}_{1(1.t)} = \begin{bmatrix} \mathbf{C}_{mm} & \mathbf{C}_{mf} \\ \mathbf{C}_{fm} & \mathbf{C}_{ff} \end{bmatrix}$$

and $\mathbf{C}_{1(1.t)}^T\mathbf{C}_{1(1.t)}$ can be written as,

$$\begin{bmatrix} \mathbf{C}_{mm}^T\mathbf{C}_{mm} + \mathbf{C}_{fm}^T\mathbf{C}_{fm} & \mathbf{C}_{mm}^T\mathbf{C}_{mf} + \mathbf{C}_{fm}^T\mathbf{C}_{ff} \\ \mathbf{C}_{mf}^T\mathbf{C}_{mm} + \mathbf{C}_{ff}^T\mathbf{C}_{fm} & \mathbf{C}_{mf}^T\mathbf{C}_{mf} + \mathbf{C}_{ff}^T\mathbf{C}_{ff} \end{bmatrix}.$$

Since all elements are equal, the four blocks can be averaged separately,

$$\frac{1}{4(M+F)^2}\mathbf{1}^T\mathbf{C}_{1(1.t)}^T\mathbf{C}_{1(1.t)}\mathbf{1}$$

$$= \frac{1}{4}\left[\frac{1}{M^2}\mathbf{1}^T(\mathbf{C}_{mm}^T\mathbf{C}_{mm} + \mathbf{C}_{fm}^T\mathbf{C}_{fm})\mathbf{1}\right.$$

$$+ \frac{2}{MF}\mathbf{1}^T(\mathbf{C}_{mm}^T\mathbf{C}_{mf} + \mathbf{C}_{fm}^T\mathbf{C}_{ff})\mathbf{1}$$

$$\left. + \frac{1}{F^2}\mathbf{1}^T(\mathbf{C}_{mf}^T\mathbf{C}_{mf} + \mathbf{C}_{ff}^T\mathbf{C}_{ff})\mathbf{1}\right].$$

One reason that all elements of $\mathbf{C}_{1(1.t)}^T\mathbf{C}_{1(1.t)}$ are identical is because the rows within each $\mathbf{C}_{xy}$ are identical (Wray and Thompson 1990), thus $\mathbf{C}_{xy}$ can be represented by X identical column vectors. The i^{th} element of a column vector of $\mathbf{C}_{xy}$ is $\frac{1}{2}r_{i(xy).t}$ where $r_{i(xy).t}$ is the additive genetic long-term contribution between (the Mendelian samplings of) ancestor i of sex x and its descendants of sex y and the $\frac{1}{2}$ is the value of the Mendelian sampling of the ancestors. Thus, equation (7) follows.

Appendix 2

Derivation of $V(A_{i(x)}) = 4\sigma_{Ax.1}^2$. $V[A_{j(x)}]$ is the variance of selected indivuals born in generation 1 of sex x ($A_{j(x)}^*$) about the mean of all the selected indivuals of sex x ($\bar{A}_{j(x)}^*$).

$V[A_{j(x)}] = V[A_{j(x)}^* - \bar{A}_{j(x)}^*] = V[A_{j(x)}^*] - V[\bar{A}_{j(x)}^*]$

$V[A_{j(x)}^*] = (1 - k_x\rho^2)$ (Pearson 1903), $\rho^2 = \sigma_{I.0}^2/\sigma_{A.0}^2$ and

$V[\bar{A}_{j(x)}^*] = \frac{1}{X}[V[A_{j(x)}^*] + (X-1)\,\mathrm{Cov}\,[A_{j(x)}^*, A_{j'(x)}^*]]$

(Woolliams et al. 1993) where

$\mathrm{Cov}\,[A_{j(x)}^*, A_{j'(x)}^*] = P_{D(x)}^* C_{D(x)}^* + P_{H(x)}^* C_{H(x)}^*$.

$P_{D(x)}^*$ ($P_{H(x)}^*$) is the probability of two selected individuals being full (half) sibs calculated as the probability of coselection of a pair of sibs (Mendell and Elston 1974) multiplied by the number of sibs,

$$P_{D(x)}^* \approx (n_f - 1)\Phi\left[\frac{\rho_{D.1}\, i_x - v_x}{(1-\rho_{D.1}^2 k_x)^{1/2}}\right],$$

$$P_{H(x)}^* \approx \left(\frac{F}{M} - 1\right)n_f\Phi\left[\frac{\rho_{H.1}\, i_x - v_x}{(1-\rho_{H.1}^2 k_x)^{1/2}}\right].$$

$C_{D(x)}^*$ ($C_{H(x)}^*$) is the covariance between breeding values of full (half) sibs after selection (see Tallis 1964, p228),

$$C_{D(x)}^* = \sigma_{A.0}^2\left[a_D + \phi[v_x, v_x, \rho_{D.1}][\rho(\rho - \rho_{D.1}u_D)\right.$$

$$+ u_D(u_D - \rho_{D.1}\rho)]/(p_x P_{D(x)}^*) + \rho^2 i_x^2 + 2\rho i_x$$

$$\left.\cdot\Phi\left[\frac{-v_x(1-\rho_{D.1})}{(1-\rho_{D.1}^2)}\right][u_D v_x - (\rho + u_D)i_x]/P_{D(x)}^*\right]$$

and $\Phi(v_x, v_x, \rho_{D.1})$ is the frequency of a bivariate normal with correlation $\rho_{D.1}$ where both truncation deviates have value v_x(so $\Phi(v_x, v_x, \rho_{D.1}) = e^{-v_x^2(1-\rho_{D.1})}/[2\pi(1-\rho_{D.1}^2)^{1/2}]$). a_D is the correlation between the breeding values of generation 1 full-sibs before selection ($a_D = \frac{1}{2}$) and u_F is the correlation between the breeding value of an individual with the index value of its full-sib (before selection). For index I_H,

$$u_D = \left\{\frac{1}{2n}\left[-\beta_1 + \beta_2\left(1 - \frac{M}{F}\right) + \beta_3\frac{M}{F}\right]\right.$$

$$\left. + \frac{1}{4}\left[\beta_2\left(1 - \frac{M}{F}\right) + \beta_3\left(1 + \frac{M}{F}\right)\right]\right\}\sigma_{A.0}/\sigma_{I.0}.$$

To calculate $C_{H(x)}^*$ replace all subscripts of D for full-sibs by H for half-sibs, $a_H = \frac{1}{4}$ and

$$u_H = \left\{\frac{1}{2n}\frac{M}{F}(\beta_2 - \beta_2)\right.$$

$$\left. + \frac{1}{4}\left[\frac{M}{F}(\beta_3 - \beta_2) + \beta_3\right]\right\}\sigma_{A.0}/\sigma_{I.0}.$$

Derivation of $b_{xy.2}$ *and* $b_{xwy.t}$.

Wray and Thompson (1990) showed that $b_{xy.2}$ can be represented as $b_{xy.2} = n_x\beta_{s2.A_{i(x)}} = n_x\beta_{s2.I2}\beta_{I2.A_{i(x)}}$ where st and It are the selection scores (1 if selected, 0 otherwise) and index scores of individuals of sex y born in generation t, β represent regression coefficients and $\beta_{st.It} \approx z_y/\sigma_{I.t}$. The coefficient $\beta_{I2.A_{i(x)}} = \frac{1}{2}\tau_x$ and thus $b_{xy.2} \approx \frac{1}{2}n_x\tau_x z_y/\sigma_{I.2}$.

Further, Wray and Thompson (1990) showed that,

$$b_{xy,t}=\frac{1}{2}\left(\frac{Y}{M}b_{xm,t-1}+\frac{Y}{F}b_{xf,t-1}+\frac{M}{X}b_{xmy,t}+\frac{F}{X}b_{xfy,t}\right) \tag{A2.1}$$

where $b_{xwy,t}$ is the regression coefficient which accounts for the additional selective advantage of the ancestor to the descendant over and above the selective advantage from the ancestor to the parent of sex w in generation $t-1$.

Similar to the case for $b_{xy,2}$, $b_{xwy,3}$ can be represented as

$$b_{xwy,3}\approx n_w\beta_{s3,I3}\beta_{I3,A_{i(x)}}$$

where the regression coefficients represent regression over and above that already accounted for between x and w, i.e., after accounting for the selection of w (as well as selection of x). The covariance between grandparent of sex x and grandoffspring before accounting for selection of w is $\frac{1}{4}\tau_w V_x$ and

$$\beta_{I3,A_{i(x)}}\approx\left[\frac{1}{4}\tau_w V_x-\frac{\frac{1}{2}\tau_x V_x\frac{1}{2}\tau_w\sigma_I^2(1-\frac{1}{4}\tau_m k_m-\frac{1}{4}\tau_f k_f)}{\sigma_I^2(1-\frac{1}{4}\tau_m^2 k_m-\frac{1}{4}\tau_f^2 k_f)}k_w\right]\Bigg/\left[V_x-\frac{(\frac{1}{2}\tau_x V_x)^2}{\sigma_I^2(1-\frac{1}{4}\tau_m^2 k_m-\frac{1}{4}\tau_f^2 k_f)}k_w\right] \tag{A2.2}$$

where $V_x=V(A_{i(x)})$, the variance of the breeding value of the ancestors. This can be approximated as $\beta_{I3,A_{i(x)}}\approx\frac{1}{4}\tau_w(1-\tau_x k_w)$. Thus $b_{xwy,3}\approx\frac{1}{4}n_w\tau_w(1-\tau_x k_w)z_y/\sigma_{I,3}$. Using similar derivations and approximations it can be shown that for $t\geqslant 4$,

$$b_{xwy,t}\approx\tfrac{1}{2}n_w\tau_w\tfrac{1}{2}(1-\tau_x k)[\tfrac{1}{2}(1-k\tau)]^{t-4}\cdot\tfrac{1}{2}(1-\tau k_w)z_y/\sigma_{I,t}.$$

In this equation $\frac{1}{2}(1-\tau_x k)$ is the reduction accounting for the selection of descendants born in generation 2 averaged over males and females (hence the average k) but dependent on the sex of the ancestor. The $\frac{1}{2}(1-\tau k)$ terms are reductions due to intermediate generations of selection, averaged over both sexes of descendants; hence the term involves average k and average τ. $\frac{1}{2}\tau_w$ is the coefficient of relationship between w and y and $\frac{1}{2}(1-\tau k_w)$ is the reduction due to selection of sex w averaged over male and females in generation $t-2$.

Thus in summary we can write,

$$b_{xwy,3}\approx\frac{1}{2}n_w\tau_w\frac{1}{2}(1-\tau_x k_w)z_y/\sigma_{I,3},$$

$$b_{xwy,t}\approx\frac{1}{2}n_w\tau_w c'_x c_w c^{t-4}z_y/\sigma_{I,t},$$

where $c_w=\frac{1}{2}(1-\tau k_w)$, $c'_x=\frac{1}{2}(1-\tau_x k)$, $c=\frac{1}{2}(c_m+c_f)$. As in Woolliams et al. (1993) it will be assumed that $\sigma_{I,2}$ will be close to its equilibrium value.

Following Wolliams et al. (1993), equation (A2.1) can be represented in matrix notation by,

$$\mathbf{b}_{x,t}=D\mathbf{b}_{x,t-1}+\frac{M}{2X}(b_{xmm,t}\ b_{xmf,t})^T+\frac{F}{2X}(b_{xfm,t}\ b_{xff,t})^T,$$

where $\mathbf{b}_{x,t}=[b_{xm,t}\ b_{xf,t}]^T$ and $\mathbf{D}$ is $\begin{bmatrix}1/2 & 1/2M/F\\ 1/2F/M & 1/2\end{bmatrix}$. Matrix $\mathbf{D}$ describes the dispersion of genes through the population from generation to generation in the absence of selection and has the property of idempotency ($\mathbf{D}^2=\mathbf{D}$). By analogous derivation to that presented in Appendix 1 of Woolliams et al. (1993) it can be shown that,

$$\mathbf{b}_{x,3}=\frac{1}{2}\sigma_I^{-1}\left[i\tau_x\begin{pmatrix}M/X\\F/X\end{pmatrix}+d_x\begin{pmatrix}i_m M/X\\i_f F/X\end{pmatrix}\right]$$

$$\mathbf{b}_{x,t}=\frac{1}{2}\sigma_I^{-1}\left[i(\tau_x+d_x+c'_x dS_{t-4})\begin{pmatrix}M/X\\F/X\end{pmatrix}+d_x c'_x c^{t-4}\begin{pmatrix}i_m M/X\\i_f F/X\end{pmatrix}\right]\quad\text{for } t\geq 4$$

where $S_t=\sum_{j=0}^{t}c^j$, $d=\frac{1}{2}(\tau_m c_m+\tau_f c_f)$ and $d_x=\frac{1}{2}[\tau_m\frac{1}{2}(1-\tau_x k_m)+\tau_f\frac{1}{2}(1-\tau_x k_f)]$. Following from this we can write in summary,

$$b_{xy,2}\approx\frac{1}{2\sigma_{I,2}}\frac{Y}{X}\tau_x i_y$$

$$b_{xy,3}\approx\frac{1}{2\sigma_{I,2}}\frac{Y}{X}(i\tau_x+i_y d_x)$$

$$b_{xy,t}\approx\frac{1}{2\sigma_{I,2}}\frac{Y}{X}(i\tau_x+id_x+idc'_x S_{t-4}+i_y c'_x dc^{t-4})$$

$$b_{xy,\infty}\approx\frac{1}{2\sigma_{I,2}}\frac{Y}{X}(i\tau_x+id_x+idc'_x S_\infty)$$

where $S_\infty=(1-c)^{-1}=2(1+k\tau)^{-1}$. Substitution of expressions of d, d_x, c'_x and S_∞ into $b_{xy,\infty}$ results in equation (11).

Appendix 3. Adaptation to equations when $(F/M)n_f>M$

The adaptation of equations in this situation affects only males selected from male parents. Therefore, in generation 2, only regression coefficient $b_{mm,2}$ is affected, resulting in $b_{mm,2}\approx\frac{1}{2\sigma_{I,2}}\tau_m i'_m$. However, in subsequent generations all $b_{xy,t}$ are affected via $b_{xmm,t}$ terms [see equation (A2.1)], and it can be shown that $b_{xy,\infty}\approx\frac{1}{2\sigma_{I,2}}\frac{Y}{X}(i_{(x)}\tau_m+G_x+Gc'_x\eta_\infty)$ where i'_m is the selection intensity appropriate to p'_m, $i_{(x)}=i$ if $x=f$ and $\frac{1}{2}(i'_m+i_f)$ if $x=m$. $G=\frac{1}{2}(\varepsilon i_m+di_f)$, $G_x=\frac{1}{2}(\varepsilon_x i_m+d_x i_f)$, $\varepsilon=\frac{1}{2}(\tau_m c_m i'_m/i_m+\tau_f c_f)$ and $\varepsilon_x=\frac{1}{2}(\tau_m\frac{1}{2}(1-k_m\tau_x)i'_m/i_m+\tau_f\frac{1}{2}(1-k_f\tau_x)]$.

Changes to the variances required for equation (22) are

$$\sigma^2_{(e)mm}=\left(1-\frac{1}{M}\right)\left(\frac{T-M}{T-1}\right)+n_m(n_f-1)U_{(e)fm}+\frac{1}{p'_m}\left(\frac{1}{p'_m}-n_f\right)U'_{(e)mm}$$

$$\sigma^2_{(g)mm}=\left[\frac{1}{p'_m}\left(\frac{1}{p'_m}-n_f\right)+n_m(n_f-1)\right]U'_{(g)mm}$$

$$\sigma_{(e)mm,mf}=n_m n_f U_{(e)fm,ff}+n_m\left(\frac{1}{p'_m}-n_f\right)U'_{(s)mm,mf}$$

$$\sigma_{(g)mm,mf}=\frac{n_m}{p'_m}U'_{(g)mm,mf}$$

where the U terms represent the Mendell and Elston (1974) probabilities of coselection of sibs, the ' implying the use of p'_m, i'_m, k'_m, and v'_m, and the use of $\rho_f+\rho_{Fe}$, ρ_{He} or ρ_m are implied in $U_{(s)fm}$ (or $U_{(s)fm,ff}$), $U_{(s)mm}$ (or $U_{(s)mm,mf}$) of $U_{(l)mm}$ (or $U_{(l)mm,mf}$) respectively. These terms are equivalent to those in the main text if p'_m, is replaced by p_m.

Appendix 4. Example

For an example population $M = 20$, $F = 200$, $n_f = 6$, $h_0^2 = 0.4$ and $\sigma_c^2/\sigma_p^2 = 0$ with selection using I_H, then terms described in the first section of the methods are $n_m = 60$, $T = 1200$, $p_m = 0.0167$, $p_f = 0.1667$, $v_m = 2.129$, $v_f = 0.967$, $z_m = 0.041$, $z_f = 0.250$, $i_m = 2.485$, $i_f = 1.499$, $i = 1.992$, $k_m = 0.885$, $k_f = 0.797$ and $k = 0.841$. If $\sigma_{P,0}^2 = 1$, then $\sigma_{A,0}^2 = 0.4$, $\sigma_{Am,0}^2 = \sigma_{Af,0}^2 = 0.1$ and $\sigma_{Aw}^2 = 0.2$. Index parameters are $\beta_1 = 0.250$, $\beta_2 = 0.700$, $\beta_3 = 0.957$, $\sigma_{I,0}^2 = 0.226$, $\tau_m = 0.957$, $\tau_f = 0.726$ and $\tau = 0.841$. Calculation of $V[A_{j(x)}]$ described in Appendix 2 requires the following: $\rho = 0.752$, $\rho_{H,1} = 0.436$, $\rho_{D,1} = 0.779$, $P_{H(m)}^* = 0.126$, $P_{H(f)}^* = 0.367$, $P_{D(m)}^* = 0.388$, $P_{D(f)}^* = 0.610$, $u_H = 0.328$, $u_D = 0.586$, $C_{H(m)}^* = 0.007$, $C_{H(f)}^* = 0.013$, $C_{D(m)}^* = 0.058$ and $C_{D(f)}^* = 0.060$. From these $V[A_{j(m)}^*] = 0.200$, $V[A_{j(f)}^*] = 0.220$, $V[\bar{A}_{j(m)}^*] = 0.018$, $V[\bar{A}_{j(f)}^*] = 0.003$ and $V[A_{j(m)}] = 0.182$, $V[A_{j(f)}] = 0.217$. Then, since $\sigma_{Ax,1}^2 = \frac{1}{4} V[A_{j(x)}]$, $\sigma_{Am,1}^2 = 0.046$, $\sigma_{Af,1}^2 = 0.054$ and $\sigma_{A,2}^2 = 0.300$. From equation (1) $\sigma_{I,2}^2 = 0.152$ and from equation (11) $b_{mm,2} = 3.05$, $b_{mf,2} = 18.41$, $b_{fm,2} = 0.23$, $b_{ff,2} = 1.40$, $b_{mm,\infty} = 2.69$, $b_{mf,\infty} = 26.85$, $b_{fm,\infty} = 0.23$, $b_{ff,\infty} = 2.34$, B_{xy} are reported in Table 3. For offspring born in generation 2, correlations are: $\rho_{H,2} = 0.308$, $\rho_{D,2} = 0.671$, $\rho_{m,2} = 0.275$, $\rho_{f,2} = 0.188$ from equations (3)–(5). Variances and covariances of contributions from parents of sex x which are not attributable to the selective advantage conferred by the parent of sex x are: $\sigma_{(e)mm}^2 = 2.45$, $\sigma_{(e)mf}^2 = 28.52$, $\sigma_{(e)mm,mf}^2 = 4.37$, $\sigma_{(e)fm}^2 = 0.21$, $\sigma_{(e)ff}^2 = 2.20$ and $\sigma_{(e)fm,ff}^2 = 0.30$ [equations (16) and (18)]. Variances and covariances of contributions from parent to offspring which are attributable to the selective advantage of the parent are: $\sigma_{(g)mm}^2 = 2.84$, $\sigma_{(g)mf}^2 = 68.52$, $\sigma_{(g)mm,mf}^2 = 13.04$, $\sigma_{(g)fm}^2 = 0.01$, $\sigma_{(g)ff}^2 = 0.38$ and $\sigma_{(g)fm,ff}^2 = 0.08$ [equations (16) and (18)]. These variances and covariances when summed, e.g., $\sigma_{(e)xy}^2 + \sigma_{(g)xy}^2$, give the one generation variances of family size presented in Table 2. The rate of inbreeding as evaluated by equation (22) is $\Delta F = 0.0254$ and with correction [equation (23)] $\Delta F = 0.0267$. In this example $M < n_m$ and therefore the adaptations presented in Appendix 3 apply. Thus, $p_m' = 0.0269$, $i_m' = 2.307$, $i_{(m)} = 1.903$, $i_{(f)} = 1.992$, $v_m' = 1.928$, $k_m' = 0.875$. For $b_{xy,t}$, $G_m = 0.077$, $G_f = 0.156$, $G = 0.116$, $\varepsilon_m = 0.155$, $\varepsilon_f = 0.315$ and $\varepsilon = 0.235$ resulting in $b_{mm,2} = 2.83$, $b_{mm,\infty} = 2.57$, $b_{mf,\infty} = 25.71$, $b_{fm,\infty} = 0.23$, $b_{ff,\infty} = 2.33$. Variances and covariances of contributions, which differ from those above, are $\sigma_{(e)mm}^2 = 2.21$, $\sigma_{(e)mm,mf} = 3.40$, $\sigma_{(g)mm}^2 = 3.47$ and $\sigma_{(g)mm,mf}^2 = 11.90$. Finally, $\Delta F = 0.0233$ and with correction (23) $\Delta F = 0.0244$.

Acknowledgements. The authors gratefully acknowledge funding from the Australian Wool Corporation (N.R.W.) and the Ministry of Agriculture, Fisheries and Food, U.K. (J.A.W. and R.T.). Drs. P. Visscher and J. van Arendonk and an anonymous referee are thanked for their comments on the manuscript.

References

Felsenstein J (1989) Alan Robertson's contributions to population genetics. In: (eds) Hill WG, Mackay TFC Evolution and animal breeding. C.A.B International, Wallingford, Oxon, England, pp 3–11

Hill WG (1979) A note on effective population size with overlapping generations. Genetics 92:317–322

Latter BDH (1959) Genetic sampling in a random mating population of constant size and sex ratio. Aust J Biol Sci 12: 500–505

Mendell NC, Elston RC (1974) Multifactorial traits: genetic analysis and prediction of recurrence risk. Biometrics 30:41–57

Robertson A (1961) Inbreeding in artificial selection programmes. Genet Res 2:189–194

Tallis GM (1961) The moment generating function of the truncated multi-normal distribution. J Roy Stat Soc B 23:223–229

Vries AG de, Van der Steen HAM, De Roo G (1990) Effects of family size in selection and testing in a closed dam line of pigs. Livest Prod Sci 24:47–63

Woolliams JA (1989) Modification of MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. Anim Prod 49: 1–14

Woolliams JA, Wray NR, Thompson R (1993) Prediction of long-term contributions and inbreeding in populations undergoing mass selection. Genet Res 62:231–242

Wray NR (1989) Consequences of selection in finite populations with particular reference to closed nucleus herds of pigs. PhD thesis, University of Edinburgh

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

Wray NR, Woolliams JA, Thompson R (1990) Methods for prediction rates of inbreeding in selected populations. Theor Appl Genet 80:503–512

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Paper 12

A THEORY OF GENETIC CONTRIBUTIONS

J.A. Woolliams and R. Thompson
Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, U.K.

SUMMARY

The concept of genetic contributions is developed and used to predict the rate of inbreeding (ΔF) for populations undergoing selection using sib indices (including mass selection as a special case). The opportunity is taken to improve upon previously published results and compare with other formulations. The use of genetic contributions was extended to genetic progress (ΔG) and it is shown how $\Delta F \propto E(r^2)$ and $\Delta G \propto E(ra)$, where r is an individual's genetic contribution and a is the Mendelian sampling component of its genotype. Using this formulation an informal solution for problems involving the simultaneous control of inbreeding and progress is given. This is used to derive a weak bound for inbreeding in relation to the genetic progress obtained. A novel class of indices involving Mendelian terms are introduced. These are compared in theory to previously published indices.

GENETIC CONTRIBUTIONS

Whilst the theory is open to generalization for overlapping generations, for the purpose of this paper discrete generations of random mating will be assumed in a population propagated by M male parents and F female parents. We will define the genetic contribution of an ancestor i in generation n_1 to descendants in a later generation n_2 to be the proportion of all distinct genealogical pathways that travel from generation n_1 to n_2 that start with ancestor i. This is equivalent to the definitions of Wray and Thompson (1990) and Woolliams, Wray and Thompson (1993) but these have been divided by a factor of $M+F$ for simplicity of expressions. Thus the genetic contribution of an individual is simply a proportion lying between 0 and ½ (since the population is assumed dioecious). The "long-term contribution" is the contribution when n_2 is much bigger than n_1 and in a well-mixed population in generation n_2 the proportion of pathways leading back from an individual that start from a particular ancestor is the same for all. The long-term contribution of ancestor i of sex x will be denoted by $r_{i(x)}$.

RATE OF INBREEDING

Wray and Thompson (1990) proved that the rate of inbreeding (ΔF) with respect to a base population could be defined in terms of long-term contributions (as defined here) by $\Delta F = \frac{1}{4}\Sigma r_{i(x)}^2$ where the sum is over the $M+F$ individuals selected from the offspring of that base population. The issue of the base population will be returned to later. Since ΔF in any population is a variable we then have $E(\Delta F) = \frac{1}{4}[ME(r_{i(m)}^2) + FE(r_{i(f)}^2)]$ where the expectations are conditional upon i being selected for breeding. Estimation of $E(r_{i(x)}^2)$ involves estimating both the expectation and the variance of $r_{i(x)}$ since $E(r_{i(x)})^2 = E(r_{i(x)})^2 + Var(r_{i(x)})$.

Expected value of the long-term contribution.

It is worthwhile considering in detail the expected value of the long-term contribution since its usefulness will be seen later to extend beyond the prediction of inbreeding alone. Unlike random selection the contribution of an individual ancestor i of sex x will depend on its breeding value since this influences the chances of descendants being selected in all generations - but with diminishing importance. The steps for predicting this contribution were derived by Wray and Thompson (1990) but they obtained the solution to their

infinite array of equations for mass selection in a recursive way. Woolliams, Wray and Thompson (1993) showed that the explicit solution for this problem could be obtained by summing a geometric series, and by using this method the expected contribution of an ancestor to all generations of descendants could be obtained. Of most interest is the long-term contribution, $r_{i(x)}$, and with mass selection:

$$E(r_{i(x)}) = (2X)^{-1}[1+\tfrac{1}{2}iS_{\infty}A_{i(x)}\sigma_P^{-1}] = (2X)^{-1}[1+i(1+kh^2)^{-1}A_{i(x)}\sigma_P^{-1}] \qquad ...(1)$$

where $A_{i(x)}$ is the deviation of the true breeding value of individual i from the expected breeding value of i given that it was selected, X is the number of parents of sex x, i the average intensity of selection over both sexes $i=\frac{1}{2}(i_m+i_f)$, S_∞ is an infinite sum equal to $(1-c)^{-1}$ where c has value $\frac{1}{2}(1-kh^2)$, and k is Pearson's variance reduction coefficient. In this context h^2 is measured in generation 2 where generation 0 is the unselected base population assumed to have a phenotypic variance of 1 and additive genetic variance h_0^2.

Wray, Woolliams and Thompson (1994) extended this methodology to include half-sib indices. Here

$$E(r_{i(x)}) = (2X)^{-1}(1 + \tfrac{1}{4}iS_{\infty}(\tau_x+\tau)A_{i(x)}\sigma_I^{-1}) \qquad ...(2)$$

where τ_x is twice the regression of the index on the breeding value of the parent of sex x and $\tau=\frac{1}{2}(\tau_m+\tau_f)$, and now $c=\frac{1}{2}(1-\tau\beta)$ where β is the covariance of the index with breeding values amongst offspring of the ancestors (ie. in generation 2). The equivalence with the result for mass selection is seen by noting $\tau_m=\tau_f=1$ and $\beta=h^2$ and $\sigma_I=\sigma_P$. τ_m and τ_f are also equal for full-sib indices. An important point to note is that as well as the scaling factor X^{-1}, the expected long-term contribution is assymmetric between males and females for general index selection with a steeper slope relating $Xr_{i(x)}$ to $A_{i(x)}$ for the sex which provides the most information on each candidate. Furthermore c is no longer directly related to the genetic variance between families as it was with mass selection.

With random selection $E(r_{i(x)})$ is simply $(2X)^{-1}$ for all ancestors of sex x since their breeding value is of no consequence to their chances of proliferation. The principles used to derive the above extend to more general indices involving parental and other information (e.g. BLUP) but the form becomes more complex; however, it is possible that re-expression of indices as dealt with in a later section may allow more simple approximate forms to be derived for these cases.

Expression for rate of inbreeding

Whilst the derivation of the expected contribution is relatively straightforward the variance of the contibution includes sampling errors arising from each individual separately and the sampling covariances arising from correlations among the index values of relatives. Explicit terms for mass selection are given by Woolliams *et al.* (1993) but the expression for ΔF is less compact and more cumbersome for half-sib indices. In this more general paper we shall not give this latter expression since many of the points of principle can be shown in the context of mass selection. In this case ΔF may be expressed in the form:

$$\Delta F = \Delta F_E + i^2[S_\infty^2-1]\rho_m[(16M)^{-1}+(16F)^{-1}] + i^2[S_\infty^2-1]\rho_f(8F)^{-1}$$
$$+ i^2B_\infty(S_\infty-1)[2\rho_m+(1+MF^{-1})\rho_f][(32M)^{-1}+(32F)^{-1}] \qquad ...(3)$$

where ρ_x is the half-sib correlation among sibs with common parent of sex x, and B_∞ is $(1-\frac{1}{2}c^2)^{-1}$. ΔF_E is the rate of inbreeding assessed assuming independent generations of selection, this treats genetic covariances among sibs and between parents and offspring as if they were of environmental origin. Such methods were termed "one generation methods" by Wray, Woolliams and Thompson (1990).

ΔF_E can be calculated using a range of methods to desired accuracy: for mass selection the covariances of selection probabilities among sibs can be calculated using the product of i^2 and ρ_x however for index selection the more accurate methods derived by Mendell and Elston (1974), and developed for this application by Wray *et al* (1994), are required. These covariances can be used either in conjunction with eigenvalue methods (eg. Woolliams, 1989) or to calculate variances of family size (Hill, 1972).

The additional inbreeding from selection over and above ΔF_E arises principally from $E(r_{i(x)})^2$ rather than

$Var(r_{i(x)})$. From the form of the expected long-term contribution given in equation (1) above we might expect terms in $i^2S_\infty^2$, but part of the expected contribution arises from the first generation of selection and is already included in ΔF_E. Since in the first generation the expected contribution is related to $iA_{i(x)}$ (rather than $iS_\infty A_{i(x)}$ in the long-term) the additional terms required for ΔF are of the form $i^2(S_\infty^2-1)$. The ρ_x come directly from $\frac{1}{4}E(A_{i(x)})\sigma_P^{-2}$. The last term in equation (3) arises from an accumulation of sampling errors since the variance of their distribution depends on the breeding value of the ancestor.

The differences between equation (3) and the expression derived by Woolliams *et al* (1993) are discussed in an Appendix. Equation (3) and its analogue for index selection result in very accurate predictions of ΔF (Table 1). The results presented include those for which the highest prediction errors are encountered. Experience has shown that the problems of prediction arise not in accounting for selection *per se* but from coping with the very high correlations among indices of sibs when h^2 is very low. It is here that the methods of Mendell and Elston (1974) are essential.

Table 1. Predicted rates of inbreeding obtained from Equation (3) and its sib-index analogue for hierarchical population structures. Each female has 6 offspring of each sex and selection was either on phenotype or a sib index (including both half- and full-sibs).

	Heritability			
	0	0.1	0.6	0.99
M=F=20				
Mass	114 (117)	146 (144)	194 (196)	166 (162)
Index	518 (503)	555 (519)	313 (305)	168 (162)
M=20, F=200				
Mass	67 (68)	87 (87)	112 (119)	88 (87)
Index	464 (452)	382 (368)	175 (180)	89 (88)

General

Caballero and Hill (1992) have suggested that the applicability of Wray and Thompson's definition was limited to random mating. Further investigation (Woolliams, unpublished) has shown that a more complete definition of the definition replaces the ¼ by ¼(1-α) where α is the extent of non-random mating as defined by Kimura and Crow (1963).

Santiago and Caballero (1994) developed an alternative derivation of ΔF for mass selection based on an argument involving drift variance. Their expression is:

$$\Delta F = \Delta F_E + i^2[S_\infty^2-1]\rho_m[(8M)^{-1}+(8F)^{-1}] + i^2[S_\infty^2-1]\rho_f(4F)^{-1} \quad \ldots(4)$$

The general form is very similar to equation (3) but the first two terms additional to ΔF_E are doubled and the third is omitted. A key difference is that equation (4) uses equilibrium parameters to derive S_∞ and ρ_x whereas equation (3) uses parameters from generation 2 (where generation 0 is the unselected, unrelated base). Equation (4) provides estimates of equal precision to equation (3) for mass selection. It is unclear as yet how these two derivations leading to (3) and (4) differ in their assumptions and approximations but it would be informative for this to be further clarified, possibly using an inductive proof.

Finally this returns to the issue of the base population. Both Wray and Thompson (1990) and Santiago and Caballero (1994) compare their predictions of inbreeding to simulations that start from an unrelated base population with genetic variance h_0^2 equal to twice the Mendelian sampling variance. The proof given by Wray and Thompson (1990) relating ΔF to $E(r^2)$ does not depend on this initial variance but it does indicate that $E(r^2)$ should be calculated for the ancestors that are the offspring of the base population. For this reason,

together with the observation that approximately half of $E(r^2)$ is determined by the contribution of these ancestors to their offspring in generation 2, Woolliams *et al* (1993) used parameters appropriate to generation 2 as an approximation. In contrast, Wray and Thompson (1990), with their recursive technique, continuously updated the parameters; whilst Santiago and Caballero (1994), with their different approach, used equilibrium parameters. It can be shown from simulation (Woolliams, unpublished) that if the rate of inbreeding is calculated with respect to an unrelated base population which has a genetic variance equal to the equilibrium variance then as predicted by the theory ΔF is related to $E(r^2)$ as before. In this case generation 2 parameters used in predictions would equal equilibrium parameters.

GENETIC PROGRESS

It seems natural to explain inbreeding in terms of our current language used to describe genetic progress (eg. intensity and accuracy of selection), so that we may obtain a deeper understanding of their dynamic relationship. In fact insight is gained by doing the reverse: expressing genetic progress in the language described above for inbreeding.

The definition of genetic progress in the context of contributions was implicit in the definition of genetic contribution, given by Wray and Thompson (1990), as the relationship of an ancestor's Mendelian sampling term with his descendants. The breeding value of an individual can be decomposed into its own Mendelian component plus the average of the breeding values of its parents; in turn, the breeding values of its parents can be decomposed into their Mendelian components plus the average of the breeding values of its grandparents; this process can be applied to the grandparents, and recursively back through the pedigree. Thus an individual's breeding value is the weighted sum of the Mendelian components of all its ancestors. A single round of selection can then be viewed as, simultaneously making an initial selection among the current generation's Mendelian components, making a second selection among the parental generation and so on. A selection alters the population mean through changing the contribution of an individual's Mendelian sampling component present in the population. Thus over many generations we have with full generality (Woolliams, in preparation):

$$E(\Delta G) = T_m E(r_{i(m)} a_{i(m)}) + T_f E(r_{i(f)} a_{i(f)}) \qquad ...(5)$$

where T_m, T_f are the total number of males and females measured with a view to selection, $r_{i(x)}$ their long-term contribution and $a_{i(x)}$ their Mendelian sampling term. The expectations are not conditional upon the selection of individual i.

There are several points that are worth emphasising about this definition. Firstly it re-affirms what our objective in selective breeding should be: it is to identify the novel variation an individual is bringing into the population and to promote or suppress its contribution accordingly. This moves away from an emphasis on selection intensity and accuracy with their associated effects on genetic variance.

Secondly, although this definition is for the long-term, similar definitions can be derived for limited time-horizons using appropriately truncated contributions $r_{i(x),t}$ (where $r_{i(x)} = r_{i(x),t}$ as $t \rightarrow \infty$). If it is assumed that new mutational variance is appearing at a rate to balance loss through inbreeding then $E(a_{i(x)}^2)=\frac{1}{2}h_0^2$ in all generations. However, if there is an effect on inbreeding on the Mendelian sampling variance then expressions for ΔG resulting from the definition given above can be simply modified to account for this, and selection limits, in the classical form discussed by Robertson (1960), can be derived.

With the infinite loci model the Mendelian sampling component for an individual is unique to itself and independent of all others. Thus, the decomposition of genetic gain into these components gives opportunities to address other problems e.g. variance of ΔG (Woolliams and Meuwissen, in preparation). Finally, whilst considering the rate of inbreeding with respect to a base generation the importance of generations in determining ΔF decreased as they moved forwards away from the base, with genetic progress the importance of generations decrease as they move back from the current generations.

As an example of the probity of the definition, the expected genetic gain in mass selection may be

considered. Firstly from equation (1), $E(r_{i(x)}a_{i(x)}|\ i$ selelected$)=(8X)^{-1}iS_-(1-k_-h^2)h_0^2\sigma_P^{-1}$, and $E(r_{i(x)}a_{i(x)}|\ i$ not selelected$)=0$. To complete the unconditional expectation note that $E(a_{i(x)}|\ i$ selected$)=\frac{1}{2}ih_0^2\sigma_P^{-1}$ and $E(r_{i(x)}|\ i$ selected$)=(2X)^{-1}$ and $Prob(i$ selected$)=XT_x^{-1}$; and, after summing over the sexes:

$$E(\Delta G) = \tfrac{1}{2}ih_0^2\sigma_P^{-1}(1+\tfrac{1}{2}S_-(1-kh^2)) = ih_0^2\sigma_P^{-1}(1+kh^2)^{-1} \qquad ...(6)$$

Here h^2 maybe that appropriate for generation 1 or 2, but $\frac{1}{2}h_0^2$ is the variance of the Mendelian sampling term. If equation (6) is equated to $ih_e^2\sigma_P$ where $h_e^2$is the equilibrium heritability then good approximations for h_e^2 are obtained. Furthermore if we replace h^2 by h_e^2 in equation (6) and then as before equate it to $ih_e^2\sigma_P$ and solve for h_e^2, a quadratic equation is obtained that is identical to that of Bulmer (1980).

COMBINING GENETIC PROGRESS AND INBREEDING

Since individuals that are not selected can make no long-term contribution the definition of inbreeding can be modified from the equation given in the first section and the following pair of equations are obtained for the rate of mean gain and the rate of loss of genetic variance in the population.

$$E(\Delta F) = \tfrac{1}{4}[T_m E(r_{i(m)}^2) + T_f E(r_{i(f)}^2)]$$

$$E(\Delta G) = T_m E(r_{i(m)}a_{i(m)}) + T_f E(r_{i(f)}a_{i(f)})$$

The expectations for both $E(r_{i(x)}^2)$ and $E(r_{i(x)}a_{i(x)})$ are now all unconditional.

The forms given above make their joint consideration more straightforward and indicate that consideration on the observed scales for both ΔF and ΔG has some justification. In the following we shall outline an informal solution to the problem of minimizing $E(\Delta F)$ for a given expected rate of genetic progress say G_{req} in a population with M male and F female parents. This problem is equivalent to maximizing the rate of gain for a given rate of inbreeding or a linear combination of the two. Further the solution is applicable to maximizing gain (may be discounted) over a given horizon with costs attached to inbreeding depression (such as used by Wray and Goddard (1994)) since for small ΔF the expression can be approximated by one linear in ΔF plus second order terms of $O(\Delta F^2)$. We shall assume that either loss of variance through inbreeding is balanced by mutational variance, or that the problem is expressed in terms of some initial (possibly 'equilibrium') parameters and that the effects of inbreeding on the mutational variance are to be ignored. For simplicity of expressions we shall also assume $T_m=T_f=T$.

To solve the problem the introduction of a Lagrangian multiplier λ is required and we need to minimize the sum over males and females of:

$$\sum_{x=m,f} TE(\tfrac{1}{4}r_{i(x)}^2 - \lambda r_{i(x)}a_{i(x)}) = \tfrac{1}{4}T\sum_{x=m,f} E(r_{i(x)}(r_{i(x)} - 4\lambda a_{i(x)})) \qquad ...(7)$$

We might anticipate that a certain amount of progress might be possible by the assortment of Mendelian sampling terms in the initial selection of the M males and F females. Let this be denoted G_w, then $G_w=\frac{1}{2}(G_{w(m)}+G_{w(f)})$ where $G_{w(x)}=i_x\phi_x\sqrt{(\frac{1}{2}h_0^2)}$ and ϕ_x is the accuracy of evaluation of $a_{i(x)}$. Assume $G_{req}>G_w$, so to obtain further progress $r_{i(x)}$ must be of the form $(2X)^{-1}+\alpha_x(a_{i(x)}-G_{w(x)})+\varepsilon_{i(x)}$ if $i_{(x)}$ is selected ; $r_{i(x)}=0$ if $i(x)$ is not selected. Therefore the minimization of (7) is equivalently for i selected:

$$\sum_{x=m,f} XE([(2X)^{-1}+\alpha_x(a_{i(x)}-G_{w(x)})][(2X)^{-1}+(\alpha_x-4\lambda)(a_{i(x)}-G_{w(x)})]-4\lambda G_{w(x)}) \qquad ...(8)$$

The minimization of equation (8) occurs when $G_{w(x)}$ is as large as possible (for $\lambda>0$) and $\alpha_x=2\lambda$. The value of λ can be obtained from equation (8), $\lambda=\frac{1}{2}(G_{req}-G_w)[E(Ma_{i(m)}^2+Fa_{j(f)}^2|\ i,j$ selected$)]^{-1}$. Finally a lower bound for ΔF can be obtained can be obtained using $(\frac{1}{4}E(r_{i(x)})^2 + (32M)^{-1} - (32F)^{-1})$ where the last two terms form the minimum error variance of a contribution in a hierarchical breeding scheme (derived from Gowe, Robertson and Latter, 1959). Therefore:

$$\Delta F \geq 3(32M)^{-1}+(32F)^{-1}+\tfrac{1}{4}(G_{req}-G_w)^2[E(Ma_{i(m)}^2+Fa_{j(f)}^2\ |i,j\ selected\ )]^{-1} \qquad ...(9)$$

The lower bound given above is not necessarily achievable, and takes no account of the fact that the

variance of the contributions will depend on α_x amongst other factors. Whilst improvements to the bound may well be possible the result does suggest that ΔF with selection will necessarily increase quadratically as ΔG increases over and above the maximum achievable from a single selection based upon the estimated Mendelian sampling terms.

Errors are inevitable and indeed the difference between ΔF for a breeding scheme and the lower bound gives some indication of the magnitude of their variance. Systems for controlling errors will be important in practical application: such systems might include factorial mating (Woolliams, 1989), family assortment (Santiago and Caballero, 1994) family restriction and a range of computer techniques. However, in all these the formulation given above is important in providing the information as to what is or is not an error and therefore undesirable! Further, it is technically possible for computational techniques to predict future pedigree development from the current position and then optimize the long-term position rather than simply optimizing one-generation on.

The lower bound may be achievable in special cases: for example if $h_0^2=1$ then the Mendelian terms will be known without error and it may be feasible to achieve the optimum value of the expected long-term contribution by selecting on $a_{i(x)}$ and allowing offspring numbers to vary with $a_{i(x)}$, analogous to the suggestion of Toro and Neito (1984). The advisability of such a policy is much less clear when $a_{i(x)}$ is only known with errors.

It is interesting to compare the problem of selection with that of clonal propagation solved by Lindgren and Matheson (1986) and Bondesson(1989). They showed that for n known genotypes $A_1,...,A_n$ used in proportions $p_1,...p_n$ with $(\Sigma p_i=1)$, to maximize clonal value $(\Sigma p_i A_i)$ with a constraint on diversity (Σp_i^2) that p_i should either be zero or linearly related to A_i. The problems of clonal propagation and selection progress are essentially identical with the two distinctions: proportional representation in one crop is replaced by the long-term contribution; full breeding values are replaced by the Mendelian sampling component.

INDEX REPRESENTATION OF THE THEORY

The foregoing theory suggests that indices should pay attention to the Mendelian sampling components of an individual breeding value. Indeed a predicted breeding value $\hat{A}_i$ can be decomposed into the estimates of the Mendelian components of itself and its ancestors: let $\hat{a}_i$ be the prediction for individual i then:

$$\hat{A}_i = \hat{a}_i + \sum_t \sum_{j=1}^{2^t} \hat{a}_{ij,t} 2^{-t}$$

where $\hat{a}_{ij,t}$ is the prediction of the Mendelian sampling term for the jth ancestor t generations back.

The expression of the predicted breeding value in this form can be used to explain the relative success of previous indices based on BLUP in reducing inbreeding with little loss in gain. Villaneuva, Woolliams and Simm (1994), Luo, Woolliams and Thompson (1994) show that indices of the form proposed by Grundy and Hill (1993; increasing the h^2 used in evaluation) were relatively more successful than the form proposed by Verrier, Colleau and Foulley (1993; partial subtraction of the estimates of the breeding values of the sire and dam). To achieve a reduction in ΔF some sacrifice in accuracy is required. In the former, a first approximation (unpublished) is of the form:

$$I \propto \hat{a}_i + \sum_t \sum_{j=1}^{2^t} \hat{a}_{ij,t} 2^{-t} c^t$$

where $c < 1$. Thus loss in accuracy arises progressively as we move back through the pedigree and is truly a discounted gene flow. The form of Verrier *et al.* (assuming for simplicity the same proportion of male and female pedigree is removed) is:

$$I \propto \hat{a}_i + \sum_t \sum_{j=1}^{2^t} \hat{a}_{ij,t} (\tfrac{1}{2}-\lambda) 2^{-t+1}$$

which, in contrast, makes no distinction over ancestral generations.

It therefore seems sensible to offer the following proposition as a flexible approach to achieve the optimum index for selection: that for a given number of parents the index should consist of the Mendelian sampling estimate plus just so many of the ancestral Mendelian sampling components that are required to achieve a desired relationship between the expectation of $r_{i(x)}$ and $a_{i(x)}$, and all others should be ignored. ie.

$$I \propto \hat{a}_i + \sum_{t} \sum_{j=1}^{2^t} \hat{a}_{ij,t} 2^{-t} c_{j,t}$$

where $c_{j,t}=1$ for $t<t^*$, $0 \le c_{j,t} \le 1$ for $t=t^*$ and $c_{j,t}=0$ otherwise. The relationship between $r_{i(x)}$ and $a_{i(x)}$ will be determined by the required genetic gain. The Mendelian terms should be peeled, working back from the candidate through its parents and grandparents etc. This has an inductive logic: if the required gain is low then it would be desirable to use only the Mendelian estimate of the candidate in order to minimize inbreeding; if this were not sufficient then an element of pedigree information is required, either on the dam or the sire; just as the Mendelian term of the candidate was the first choice so one of the Mendelian terms of the parents would be the next. At the next stage the choice would lie between the other parent's Mendelian term and the Mendelian terms of the previously chosen parent's parents. This process continues recursively, at each stage there are a set of Mendelian terms to be compared. Each term chosen would be used in full except for the last term which may only be partly needed. At each stage choices are needed since one ancestral component may be more beneficial (or perhaps more credible) than another. Finally to minimize the variance of the contributions, and therefore to ensure a truly optimal solution a further optimization is required over differing numbers of parents.

There have been implicit suggestions that true BLUP evaluation should be avoided (Quinton, Smith and Goddard, 1992; Colleau, 1992; Grundy *et al*, 1993). This is because we do not know how to use the results to best advantage in selection schemes. The representation given above re-emphasises the need for BLUP, not to estimate the entire breeding value, but instead to estimate an individual's Mendelian term which is simply a linear transformation of the breeding values that are currently estimated.

CONCLUSIONS

The concept of contributions has not only led to an understanding of the process of inbreeding, but also to genetic progress: whereas inbreeding can be related as the expected squared contribution, the genetic progress can be similarly related to the covariance of the contributions with Mendelian sampling components. Thus, inbreeding and genetic progress, intrinsic to breeding schemes, are now capable of description by properties of the elements of the relationship matrix. Conversely we can see how the form of the relationship matrix in the long-term is determined by the characteristics and performance of the breeding scheme and we have the opportunity, by our selection and breeding decisions, to influence its development in order to achieve a desired goal.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding by the Ministry of Agriculture, Fisheries and Food (UK), the Milk Marketing Board of England and Wales and the Meat and Livestock Commission.

REFERENCES

BONDESSON, L. (1989) Report 27, Swedish University of Agricultural Sciences, Umeå.
BULMER, M.G. (1980) Mathematical theory of quantitative genetics. Clarendon, Oxford.
CABALLERO, A. and HILL, W.G. (1992) Genetics 130:909-916
COLLEAU, J.J. (1992) Genetics, Selection, Evolution 24:345-361
GOWE, R.S., ROBERTSON, A. and LATTER, B.D.H. (1959) Poultry Science 38:462-471
GRUNDY, B. and HILL, W.G. (1993) Animal Production 56:427A
HILL, W.G. (1972) Theoretical Population Biology 3:278-289
KIMURA, M. and CROW, J.F. (1963) Evolution 17:279-288
LINDGREN, D. and MATHESON, A.C. (1989) Silvae Genetica 35:173-177
LUO, Z.W., WOOLLIAMS, J.A. and THOMPSON, J.A. (1994) Animal Production (submitted)
MENDELL, N.C. and ELSTON, R.C. (1974) Biometrics 30:41-57
QUINTON, M., SMITH, C. and GODDARD, M.E. (1992) Journal of Animal Science 70:1060-1067
ROBERTSON, A. (1960) Proceedings of the Royal Society 153:234-239
TORO, M.A.and NEITO, B. (1984) Genetical Research 44:347-349
VERRIER, E., COLLEAU, J.J. and FOULLEY, J.L. (1993) Theoretical and Applied Genetics (in press)
VILLANUEVA, B., WOOLLIAMS, J.A. and SIMM, G. (1994) Genetics, Selection and Evolution (submitted)
WOOLLIAMS, J.A. (1989) Animal Production 49:1-14
WOOLLIAMS, J.A., WRAY, N.R. and THOMPSON, R. (1993) Genetical Research 62:231-242
WRAY, N.R. and THOMPSON, R. (1990) Genetical Research 55:41-54
WRAY, N.R., WOOLLIAMS, J.A. and THOMPSON, R. (1990) Theoretical and Applied Genetics 80:503-512
WRAY, N.R., WOOLLIAMS, J.A. and THOMPSON, R. (1994) Theoretical and Applied Genetics (in press)
WRAY, N.R. and GODDARD, M.E. (1994) Genetics, Selection and Evolution (in press)

APPENDIX

Equation (3) differs from Woolliams *et al.* (1993) in two aspects. Firstly, the derivation in that paper implicitly assumed that the expected long-term contribution had a term quadratic in $A_{i(x)}^2$. However, as noted in that paper the sum of all contributions from a single sex must always equal ½. This constraint was invoked to justify the replacement of $A_{i(x)}$ by $(A_{i(x)}-A_{bar})$ throughout (where A_{bar} is the mean breeding value of the individuals of sex x that are actually selected, which affects the calculation of ρ_x), but was not applied to the quadratic term. This is an omission. When applied it is seen that the quadratic term can make no contribution to inbreeding of $O(A_{i(x)}^2)$. Consequently all terms arising from this quadratic effect disappear. This results in considerable simplification. Secondly, the underlying theory of the approach is to predict the proliferation of the geneological pathways in the pedigree of the population: for any given ancestor the chances of a particular pathway being extended is not only dependent on its own genotype but also on that of its mate. For example when $M=F$ and full-sib families are produced then the proliferation of the ancestor's and its mate's pathways are inextricably linked. This argument suggests that when considering ΔF, long-term contributions should not only be considered conditional on the ancestors breeding value but also on those of its mates. The first of these modifications to Woolliams *et al* (1993) reduces the predictions of inbreeding, whilst the second increases them. Taken together they result in greater precision.

Paper 13

A RECIPE FOR THE DESIGN OF BREEDING SCHEMES

J. A. Woolliams

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, United Kingdom.

SUMMARY

The theory of genetic contributions is developed to provide a recipe for the design of breeding schemes to maximize gain with constraints on rates of inbreeding. It is shown that for random mating and Poisson litter sizes, assuming an equilibrium state of the breeding scheme that:

$$E[\Delta G] = \Sigma X_k E[\mu_k a_k] \quad (1)$$

$$E[\Delta F] = \tfrac{1}{2}\Sigma X_k E[\mu_k^2] \quad (2)$$

where the μ are expected long term contributions conditional upon the complete set of variables conferring selective advantage in the offspring. The sums are over breeding categories defined by sex, age and purpose. The expected long-term contribution can be derived from knowledge of the gene flow and two regression models describing: (i) the selection score of the parent regressed upon the variables conferring advantage; and (ii) the variables conferring advantage in the selected offspring regressed on the same variables in the parent.

INTRODUCTION

Woolliams and Thompson (1994) put forward the theory of genetic contributions which unified in a general way the rates of genetic gain and inbreeding in breeding schemes. The equations presented were:

$$E[\Delta G] = T_m E[r_m a_m] + T_f E[r_f a_f] \quad (3)$$

$$E[\Delta F] = \tfrac{1}{4}(T_m E[r_m^2] + T_f E[r_f^2]) \quad (4)$$

Here ΔG and ΔF are rates of gain and inbreeding respectively; T, the number of candidates; r, the long-term genetic contributions; a, Mendelian sampling terms; and subscripts m and f denote the sex. The expectations are unconditional. These equations gave an approach towards the estimation of rates of gain and inbreeding, in particular the latter, where alternative options were limited. They showed that for mass selection in discrete generations, the approach yielded answers for equilibrium rates of gain that were identical to those derived from Bulmer (1980). The approach had been used to predict rates of inbreeding in selected populations with some success in the special cases of mass selection and sib indices.

However a problem with equations (3) and (4) is that they are given in terms of the *observed* long-term contributions and so their use requires distributional models for r. The expectation is required for estimating genetic gain and in addition the variance is required for estimating inbreeding rates. This abstract describes an approach to the problem that results in a straightforward estimation process for both rates of gain and inbreeding which is based upon two tractable regression models.

METHODS

The key approach taken is to assume an equilibrium state in the population. In this state the parameters of the distribution of long-term contributions for selected individuals in one generation are identical to the parameters in the next.

Expected long-term contributions. To obtain estimates of genetic gain, and as a first step to obtaining estimates of rates of inbreeding, the expected long-term contribution for a selected parent needs to be derived as a function of terms which are related to the Mendelian sampling term of the individual. For mass selection and sib indices it is sufficient to consider only the breeding value as a whole, but for selection methods utilizing BLUP the terms should include both the estimated breeding value and the prediction error.

The use of equilibrium can be illustrated best by returning to an analogy used by Wray and Thompson (1990). Suppose a random number, n, of random variables s occurs subsequent and independent to the decision on n. Then if $v = \sum s$, the expectation $E[v] = \mu_n \mu_s$. Let the expectations of v and n be linearly related to a variable A, so that $E[v] = \alpha + \beta A$, $E[n] = 1 + \phi A$; and that the variables s have the same distributional property as v, that is $E[s] = \alpha + \beta A^+$, but their indexing variable A^+ is different from A and has a regression relationship with A given by πA plus a random error. Then equating terms in A resulting from $E[v] = \mu_n \mu_s$, gives a relationship between α and β, namely:

$$\beta = (1 - \pi)^{-1} \phi \alpha \qquad (5)$$

To make the analogy more direct, note that the long-term contribution of individual i is given by:

$$r_i = \tfrac{1}{2}\left(\sum r_{j(m)} + \sum r_{j(f)}\right) \qquad (6)$$

where the sums are taken over its male and female offspring, and since unselected offspring have no long-term contribution these sums may be restricted to the selected offspring. For mass selection (as an example) where the long-term contribution can be described as a regression on breeding value: if we are in an equilibrium then the relationship of the long-term contribution of i with its own breeding value is identical in form to the relationship of the long-term contribution of the offspring to their own breeding value. Furthermore the breeding value of the offspring can be described as a linear regression on the parental breeding value. Therefore the analogy is between v and the long-term contribution of the parent, s and the long-term contribution of a selected offspring, and between n and the numbers of male and female offspring selected. The remaining task is to generalize the result to the two sexes, to several variables describing selective advantage, and to develop the regression equations required.

There are two regression models required. These are: (ϕ), the regressions of selection success as a parent on the variable(s) conferring selective advantage to the offspring; (π), the regressions of the variable(s) conferring selective advantage in the selected offspring on the variable(s) in the parent.

Extension to overlapping generations. The methodology described above can be extended beyond discrete generations to overlapping generations. In this context, the concept of gene glow

between groups of individuals differing in sex or age or breeding purpose is needed (Hill, 1974). However the development of Hill does not account for the inheritance of selection advantage. In particular it assumes that the selection from a cohort is independent of the breeding group of the candidate's parent. This is particularly a problem with pre-determined generational structures, since genetic progress will often give offspring of younger age groups a selective advantage over other candidates (this will depend on the selection procedures but will occur in mass selection). A further consequence of this selective advantage is that the origin of the parents of a selected group will be dependent on the degree of selection taking place. As groups grow older the selection pressures change, so the origin of the parents of the selected individuals will also change with age. Therefore gene flow with selection is an *n*th order process where *n* is the maximum age of any parental group, in contrast to the first order process described by Hill (1974).

If the standard gene flow matrix for a breeding structure is denoted by G_0, elements g_{kl} represent the proportions of genes currently in category k that arise from category l in the previous time period. To obtain the equilibrium long-term contributions a modified matrix is required (G) in which each row represents a category of *selected* individuals, and with the elements of each row representing the transfer of genes through breeding from the different categories l. First approximations of these elements can be based upon G_0 and the regressions for success score developed by Wray and Thompson (1990) and Wray *et al.* (1994).

Solutions. Let $\mu_{i(k)}$ be the expected long-term contribution of category k and n_k be the number of selected individuals for that category, $\mu_{i(k)} = (\alpha_k + \beta_k A_i)/n_k$, where α and β are the vectors with elements α_k and β_k for each category k. The vector α is a right eigenvector of G^T with a scaling factor such that the sum of the elements total L^{-1}, where L is the generation interval. The generation interval is the interval over which the sum of all long-term contributions sum to 1. The average age of the parents at the birth of the next cohort is only a first approximation to this. The two are equivalent when there is no selection or when each sex is bred at a single age. The distinction is simply a logical extension to what is already practiced: offspring that are not bred as potential replacements are excluded from consideration in the 'average age of parents'. The definition implied here weights the ages of those that are selected by the chances that they will produce competitive offspring, rather than offspring which *a priori* have little chance of success.

The vector β can be derived from a matrix extension of the initial analogy (equation 5):

$$\beta = (I - G^T \otimes \Pi^T)^{-1} (G^T \otimes \Phi^T) \alpha$$

where $\otimes$ is element by element multiplication of the matrices. The matrix Π has elements π_{kl} which are the regression coefficients of a variable determining selective advantage of a *selected* individual in category k on a parent of category l. The matrix Φ has elements ϕ_{kl} which are the regression coefficients for selection score in category k of a parent of category l. These require further generalization for multiple variables conferring advantage.

For discrete generations, G = (½ , ½ | ½ , ½) in all cases. Substitution for G, and using regression models from results of Woolliams *et al.* (1993) for mass selection gives previously

published results. Use of the formulae for deriving ΔG give identical results to Rendel and Robertson (1950) when both sexes have single categories, but marginally different results for more complex problems. The advantage of the method presented is the ability to identify the expected development of individual pedigrees (i.e. between and within groups), and consequently use these to derive the rate of inbreeding.

Rates of inbreeding. The spirit of the derivation can be seen from the analogy given previously. For v, $\sigma_v^2 = \mu_n \sigma_s^2 + \mu_s^2 \sigma_n^2$: (i) if n has a Poisson distributio, $\mu_n = \sigma_n^2$; (ii) equilibrium conditions are imposed so that $E[v^2] = E[s^2]$; then an expression for the variance of v can be obtained:

$$\sigma_v^2 = (1-\mu_n)^{-1} \mu_n \mu_v^2$$

The principles applied to the analogy can be applied to long-term contributions using equation (6). The resulting expression is straightforward: with Poisson litter sizes and random mating the rate of inbreeding per cohort is one half of the sum of squared lifetime expected long-term contributions for selected individuals in that cohort:

$$E[\Delta F] = \tfrac{1}{2} \Sigma X_k E[\mu_k^2]$$

where the sum is over all possible categories k and X_k is the number of parents in category k. However it is important to note that the expectation must include all the variables conferring selective advantage, such as the breeding value of mates, permanent environmental variance etc. Nevertheless the expectation is tractable, derived from the two regression models of π_{ij} and ϕ_{ij}. In the simplest case of random selection the μ are simply $(2M)^{-1}$ and $(2F)^{-1}$ and application of equation (2) yields Wrights formula $(8M)^{-1} + (8F)^{-1}$. Application to mass selection yields the result, but *not* the principle of Santiago and Caballero (1995), and when applied to sib indices simplifies the results of Wray *et al.* (1994).

CONCLUSIONS

Equations (1) and (2) provide a general and tractable means of designing complex breeding schemes, incorporating overlapping generations and a variety of inheritance models, such as additive maternal variance, and selection methods.

ACKNOWLEDGEMENTS

The author is grateful for financial support from the Ministry of Agriculture, Fisheries and Food.

REFERENCES

Bulmer, M.G. (1980) "Mathematical theory of quantitative genetics". Clarendon, Oxford.

Hill, W.G. (1974) *Anim. Prod.* **18**:117-139

Rendel, J.M. and Robertson, A. (1950) *J. Genet.* **50**:1-8

Santiago, E. and Caballero, A. (1995) *Genet.* **139**:1013-1030

Woolliams, J.A. and Thompson, R. (1994) *Proc. 5th Wld.Cong. Genet. Appl. Livest. Prod.* **19**:125- 134

Woolliams, J.A., Wray, N.R. and Thompson, R. (1993) *Genetical Research* **62**:231-242

Wray, N.R. and Thompson, R. (1990) *Genetical Research* **55**:41-54

Wray, N.R., Woolliams, J.A. and Thompson, R. (1994) *Theoret. Appl. Genet.* **87**:878-892

Paper 14

Anim. Prod. 1989, **49**: 1-14

0003-3561/89/8362000115$02·00

MODIFICATIONS TO MOET NUCLEUS BREEDING SCHEMES TO IMPROVE RATES OF GENETIC PROGRESS AND DECREASE RATES OF INBREEDING IN DAIRY CATTLE

J. A. WOOLLIAMS
AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

ABSTRACT

The effect of changes in the mating system on the rates of genetic progress and rates of inbreeding are considered for MOET nucleus breeding schemes. Methods are derived to calculate best linear unbiased predictors of breeding value in MOET schemes and the rate of inbreeding under selection. These are applied to different mating systems in which the numbers of sires and dams, and the number of offspring per sire and offspring per dam, remain constant.

Results showed that compared with nested mating systems, factorial mating systems in which the maternal half-sibs are produced instead of full-sibs, could increase genetic progress by 1·12-fold with no additional inbreeding. The increased progress arose through an increase in the selection intensity applied. The rates of inbreeding derived were found to be approximately double those estimated by the formula of Wright (1931) in the absence of selection.

In practice, even if a complete factorial system were to increase the generation interval and consequently reduce progress below that predicted, changes in the mating system avoiding this problem could be implemented that would be of immediate benefit.

INTRODUCTION

METHODS for using multiple ovulation and embryo transfer (MOET) in genetic improvement have been under development since being suggested by Land and Hill (1975). For dairy cattle, Nicholas and Smith (1983) suggested that MOET nucleus breeding schemes could offer improved rates of genetic progress over progeny testing in dairy cattle improvement. They described two possible schemes, an adult (A) scheme whereby selection of candidates takes place after one lactation in a female and a juvenile (J) scheme in which selection takes place before any lactation in the candidates. These rates were later revised by Woolliams and Smith (1988) who showed that A schemes, whilst improving upon conventionally run progeny testing schemes, could only offer progress equivalent to an efficient progeny testing system while J schemes could offer rates 1·30 times greater. Furthermore, the authors also showed that both MOET schemes, when combined with a proven juvenile predictor, were superior to progeny testing systems.

However, the calculations used in assessing these schemes with finite resources do not allow for bias due to the correlation of index values of close relations or consider penalties for inbreeding. This paper reviews the causes of these correlations and puts forward modifications to MOET schemes aimed at reducing the occurrence of high correlations, thereby improving rates of progress and/or reducing rates of inbreeding.

METHODS

Modifications to breeding structure

MOET nucleus breeding schemes were introduced as a way of creating extended families so that a candidate for selection had a range of relatives which contributed to the estimation of its breeding value. In progeny testing, evaluation of a bull is by the performance of its daughters, and of a cow

by its parentage and its own performance. In adult MOET schemes no progeny testing as such occurs and breeding value estimation comes from information on paternal half-sibs, dams, full-sibs and, in the case of a female candidate, on its own performance. Initial MOET nucleus schemes used full-sibs primarily because it is the result in practice of achieving the ideal of obtaining all eggs from a single flush — an ideal that is not always met.

A penalty arising from the production of full-sibs in index selection is the increased potential for inbreeding through the high correlation of the index amongst them. Nicholas and Smith (1983) and later authors overcame this by restricting male selection to one per full-sibship, thus reducing the selection proportion n-fold (where n is the number of male full-sibs). However, if full-sib groupings were avoided this reduction in selection intensity may not be required.

It is then natural to consider if alternative mating schemes are competitive. Each donor cow could be mated to two or more bulls in such a way that the expected number of offspring per bull and per cow would remain the same. Note that each individual would thus have fewer full-sibs but just as many maternal and paternal sibs as before, so the relative accuracy of such mating schemes needs investigation.

Therefore, modifications to MOET have been considered and compared with the standard form. The scheme considered as a base has the following parameters.

Thirty-six donor cows are mated to four bulls per generation, and each donor cow produces eight offspring assumed (with the aid of embryo sexing) to be four male and four females and each bull 72 offspring equally split between males and females. Three systems have been considered.

I. Each bull is mated to nine donor cows to produce full-sib families of size eight (4♂, 4♀) and all maternal sibs are full-sibs.

II. Each bull is mated to 18 donor cows to produce full-sib families of size four (2♂, 2♀) and each donor is mated to two bulls. With 36 donors and four bulls, six cows are mated to bulls 1 and 2, six to bulls 1 and 3, etc.

III. Each bull is mated to each cow to produce full-sib families of size two (1♂, 1♀).

Each system has been considered either unrestricted (IU, IIU, IIIU) or restricted as follows: IR only one male per full-sibship can be selected; IIR as for IR; IIIR only one male per paternal half-sibship can be selected (by design, no pairs of full-sib males are produced). System IR is that proposed by Nicholas and Smith (1983).

Assessment of genetic progress

The genetic parameters assumed for progress are a heritability for yield (h^2) of 0·25, a genetic correlation of 1 between successive lactations and a repeatability of yield over successive lactations (t) of 0·35. The generation interval (L) for adult schemes has been taken as 3 years 10 months (3·83 years), but scheme III has also been considered with L = 4 years 2 months (4·17 years) (this is to account for possible extra collections of embryos). All juvenile schemes have been considered with L = 2 years and additionally scheme III has been considered with L = 2 years 4 months (2·33 years).

The general formula for progress (phenotypic standard deviations per annum, $\triangle G$) has been taken as

$$\triangle G = \tfrac{1}{2}(i_m r_m + i_f r_f) L^{-1} h \sigma_P$$

where i denotes selection intensities, r denotes selection accuracy (i.e. correlation of selection index and breeding value), subscripts m and f denote male and female respectively and σ_P is the phenotypic standard deviation. Selection intensity has been calculated assuming (i) infinite sample size and also (ii) finite sample size with correlated index values. The derivation of r_m and r_f is shown in APPENDIX 1 for adult MOET schemes and utilizes the balance of the mating designs to derive prediction error variances of the best linear unbiased predictors. The estimation of finite selection intensities is described in the next section.

For juvenile MOET schemes the selection is on pedigree alone and the accuracy (r_J) used (the same for males and females)

assumes that at the time of selection the sire and dam of the calf are assessed to the same accuracy as for an adult scheme, i.e. $r_J = (\frac{1}{2})(r_m^2 + r_f^2)^{\frac{1}{2}}$ where r_m and r_f are the male and female accuracies derived for the adult MOET scheme with the same mating structure.

Further results on the effect of changing the value of h^2 were considered for adult schemes only. Apart from $h^2 = 0{\cdot}25$, values of 0·1, 0·5 and 0·7 were assumed. In all these t was taken to be $1{\cdot}4\ h^2$.

Assessment of inbreeding

The expected rates of inbreeding have been calculated by recurrence relations and the details are given in APPENDIX 2.

The calculation of inbreeding rates depends on the frequency of co-selection of relatives of varying degree and requires the estimation of a range of probabilities of selection of related animals, denoted by $q(..:.)$, whose definition and parameters are described in APPENDIX 2. This problem is theoretically complex. With normality assumptions some special cases are tractable. One assumes the family structure in which there are s families of size n, where families are independent of each other and consist of only one relationship amongst family members e.g. paternal half-sibs (Hill, 1976; Rawlings, 1976).

However, in MOET nucleus schemes family structures are more complex involving both half-sibs and full-sibs and, in the modifications suggested, are not discrete. To overcome this problem the values of $q(..:.)$ can be estimated by Monte Carlo techniques. For each scheme, selection indices involving a linear combination of sire, dam and individual merit can be derived that give index values approximately distributed among relatives of differing degree according to the theory derived in APPENDIX 1. This allows simple simulation of a single stage selection process and hence an estimate of the probabilities $q(..:.)$. The values of $q(..:.)$ used in calculations are the mean of 1000 independent replications. In the course of this simulation the appropriate values of the products $i_m r_m$ and $i_f r_f$ required to measure progress in the finite population were also estimated.

RESULTS

Variance and covariance of estimated breeding values

Table 1 shows the variance and covariances of the estimated breeding values for the

TABLE 1

Index accuracy and variance, and the covariance and correlations among relatives for a variety of mating systems in adult MOET nucleus schemes

System	Candidate	Index		Covariance (and correlation) with†					
		Variance	Accuracy	♀FS	♀PHS	♀MHS	♂FS	♂PHS	♂MHS
I	♀	0·1080	0·657	0·090 (0·835)	0·052 (0·483)		0·084 (0·922)	0·050 (0·551)	
	♂	0·0775	0·557	‡	‡		0·078 (1)	0·048 (0·623)	
II	♀	0·1067	0·653	0·089 (0·833)	0·051§ (0·477)	0·040 (0·372)	0·083 (0·922)	0·049§ (0·545)	0·036 (0·400)
	♂	0·0758	0·551	‡	‡	‡	0·076 (1)	0·047§ (0·616)	0·031 (0·415)
III	♀	0·1061	0·652		0·050 (0·473)	0·039 (0·363)	0·082 (0·921)	0·048 (0·540)	0·035 (0·388)
	♂	0·0750	0·548		‡	‡		0·046 (0·611)	0·030 (0·400)

† FS = full-sib, PHS = paternal half-sib, MHS = maternal half-sib.

‡ Figures the same as for female candidate with male relatives.

§ Due to the design these values vary slightly since the MHS of the individuals chosen can be either unrelated or PHS.

different mating systems in adult MOET nucleus schemes calculated by the methods described in APPENDIX 1. For system I it is straightforward to calculate these (co) variances by alternative selection index methods described by Woolliams and Smith (1988) and Woolliams (1989). The difference between these two methods in a completely balanced case with no nuisance parameters is not expected to be large differing only in the amount of information used. For example, using best linear unbiased prediction (BLUP) the paternal half-sibs of a candidate have their dam's performance taken into account for evaluating the information they provide on the candidate's sire. BLUP estimation trivially increased the selection accuracy for system I by 1·005 and 1·0025 for males and females, respectively.

Table 1 shows that the magnitude of the variances and the covariances of the index values amongst individuals of a specified relationship change, but only slightly, according to the system of mating. The accuracy of the index varies from 0·65 to 0·66 for females and 0·55 to 0·56 for males according to the mating system. The correlations of index values among relatives show only small changes: for example between a male and a female full-sib the

TABLE 2

Index accuracy and variance, and the covariance and correlations among relatives for a variety of mating systems in juvenile MOET nucleus schemes

		Index		Covariance (and correlation) with		
System	Candidate	Variance	Accuracy	FS	PHS	MHS
I	♀ or ♂	0·046	0·431	0·046 (1)	0·019 (0·418)	
II	♀ or ♂	0·046	0·427	0·046 (1)	0·019 (0·415)	0·027 (0·585)
III	♀ or ♂	0·045	0·425	0·045 (1)	0·019 (0·414)	0·026 (0·586)

TABLE 3

Rates of genetic progress (△G) in adult MOET nucleus schemes according to different mating systems (ES, example scheme; LS, large scheme)

System	L	Restrictions on selection	Selection proportion† for males	Selection intensity‡ ES		LS		$\triangle G$ (×1000) (phenotypic s.d. per annum) ES	LS
IU	3·83	None	0·028	1·801	1·162	2·295	1·271	115	137
IR		No FS	0·111	1·448	1·162	1·705	1·271	102	115
IIU	3·83	None	0·028	1·858	1·169	2·295	1·271	116	136
IIR		No FS	0·056	1·693	1·169	2·019	1·271	111	126
IIIU	3·83	None	0·028	1·884	1·170	2·295	1·271	117	136
IIIR		No PHS	1 from 36	2·118	1·170	§	1·271	97	101
IIIU	4·18	None	¶			¶		107	124
IIIR		No PHS						89	93

† Selection proportion of females was 1/4 for all systems.
‡ Intensity derived for ES by simulation and for LS from the tables of Becker (1975).
§ Selection constrained to be within sire families, thus no large sample intensity is appropriate.
¶ Unchanged by L.

correlation is 0·92 and between paternal half-sibs the correlations are close to 0·48, 0·54 and 0·61 for two females, a male and a female and two males, respectively.

Table 2 shows similar information for juvenile MOET nucleus schemes. There is no distinction between males and females since juvenile schemes rely entirely on pedigree information and have no contribution from candidate performance. All the mating systems have an index accuracy of 0·43 an unsurprising result since mating system had little effect on the accuracy of adult schemes from which r_J is directly derived. Likewise correlations among relatives were only slightly changed by mating system.

Rates of genetic progress

Estimated rates of genetic progress for adult and juvenile MOET nucleus schemes are shown in Tables 3 and 4.

For the adult scheme used as an example, mating system IR (proposed by Nicholas and Smith, 1983) has a rate of genetic progress 0·102 phenotypic s.d. per annum. With a similar restriction against full-sib selection mating systems IIR and IIIU (L = 3·87) offer 0·111 and 0·117 phenotypic s.d. per annum, respectively, 1·09- and 1·15-fold increases over system IR. With large schemes the rates of progress are 0·115, 0·126 and 0·136 phenotypic s.d. per annum for systems IR, IIR and IIIU (L = 3·87), respectively (1·10- and 1·18-fold increases over system IR). Since mating system had little effect on the accuracy, the increases result from the increases in selection intensity possible. In small schemes this is largely, but not entirely, due to selection proportion; since for the same generation interval marginally better rates are obtained with systems IIU and IIIU (compared with IU). The advantages of systems IIU and IIIU disappear completely when restrictions on size are removed and large sample selection intensities can be used. It is noteworthy that, providing the generation interval is similar, 0·96 of the response in the example adult scheme using system IR could be obtained by employing system IIIR and selecting males within sire-families.

In practice, it may not be possible to operate system III with the same generation interval as systems I and II. If the generation interval of 4·18, rather than 3·83 years, is used then system IIIU still offers faster progress than a restricted system IR but the benefit is reduced to 1·06-fold, and is less than the benefits of system IIR.

In juvenile MOET nucleus schemes the

TABLE 4

Rates of genetic progress (△G) in juvenile MOET nucleus schemes according to different mating systems (ES, example scheme; LS, large scheme)

System	L	Restrictions on selection	Selection proportion† for males	Selection intensity‡ ES		Selection intensity‡ LS		△G (×1000) (phenotypic s.d. per annum) ES	△G (×1000) (phenotypic s.d. per annum) LS
IU	2	None	0·028	1·940	1·164	2·295	1·271	167	191
IR		No FS	0·111	1·531	1·164	1·705	1·271	145	160
IIU	2	None	0·028	2·002	1·182	2·295	1·271	170	190
IIR		No FS	0·056	1·802	1·182	2·019	1·271	159	175
IIIU	2	None	0·028	2·028	1·185	2·295	1·271	171	189
IIIR		No PHS	1 from 36	2·118	1·185	§	1·271	149	153
IIIU	2·33	None	‖	‖		‖		147	163
IIIR		No PHS						128	131

† Selection proportion of females was 1/4 for all systems.
‡ Intensity derived for ES by simulation and for LS from the tables of Becker (1975).
§ Selection constrained to be within sire families, thus no large sample intensity is appropriate.
‖ Unchanged by L.

considerations are very much the same. With restrictions on male full-sib selection in the example scheme the rates of progress for systems IR, IIR and IIIU are 0·144, 0·159 and 0·172 phenotypic s.d. per annum when each is considered with a generation interval of 2 years; 1·10- and 1·19-fold increases for systems IIR and IIIU over system IR. With large sample selection intensities similar increases are observed. In juvenile schemes, using system IIIR offers faster rates of progress than using system IR providing generation intervals can be kept the same. However, in juvenile schemes, should practical considerations increase the generation interval of system III by 4 months then all its advantages over system IR described above all but disappear.

To consider further the effect of the proportion of males selected in the example scheme with the various mating systems upon rates of progress further simulation was carried out. In system IR only four out of 144 males are selected but, because of restrictions, they are chosen from among the best 16 males of the 144 available. Thus, likewise, in system IIR the four selected males come from the best eight males (four sets of pairs of full-sibs). The selection proportion can be increased progressively by selecting four males (not full-sibs) at random from the best five, six, seven and eight pairs of full-sibs, the last having an equal selection proportion with system IR. In system IIIU there are no male full-sib pairs and so the selection proportion is progressively increased simply by selecting four males at random from the best five up to 16 males, with again the last having an equal selection proportion to system IR. As before, 1000 replications were made for each proportion.

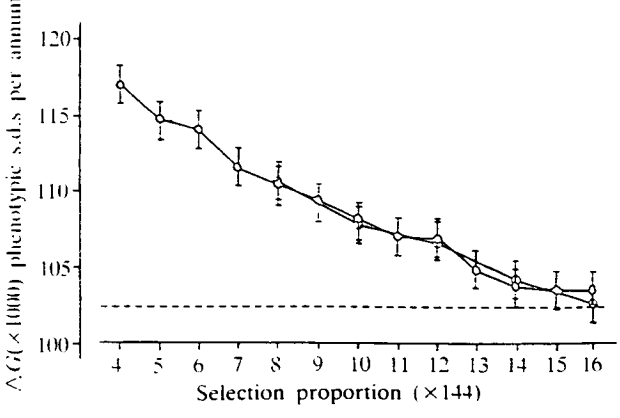


FIG. 1. The rate of genetic progress ($\triangle G$) according to selection proportion for systems IIIU (●) and IIR (○) in adult MOET nucleus schemes. The bars show ±2 s.e.s calculated from the simulation variance. Maximum $\triangle G$ for system IR is indicated by the broken line and has s.e. 0·35 phenotypic s.d.s per annum on the scale presented.

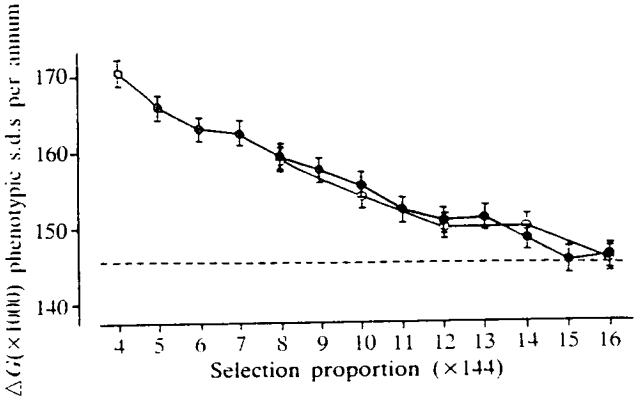


FIG. 2. The rate of genetic progress ($\triangle G$) according to selection proportion for systems IIIU (●) and IIR (○) in juvenile MOET nucleus schemes. The bars show ±2 s.e.s calculated from the simulation variance. Maximum $\triangle G$ for system IR is indicated by the broken line and has s.e. 1·05 phenotypic s.d.s per annum on the scale presented.

The results of these simulations are presented in Figures 1 and 2. The standard errors shown were derived from the variance among replications. The figures show for both adult and juvenile MOET nucleus schemes the expected reduction in progress as selection proportion is increased: at the same selection proportion there is only a marginal difference between IR, IIR and IIIU, agreeing with the previous result for unrestricted systems. System IIIU is still superior but only when considered at the same generation interval. With a longer generation interval, at the same selection proportion system IIIU only offers 0·98 and 0·91 times the progress of system IR in adult and juvenile schemes, respectively.

TABLE 5

Estimated q(.,.:.) *and effective population size* (N_e) *for a variety of mating systems in adult MOET nucleus schemes*

Mating system	Restrictions on selection	q(m.m:.)			q(m.f:.)			q(f.f:.)†			N_e‡		Inbreeding rate§ (% per annum)	
		FS	PHS	MHS	FS	PHS	MHS	FS	PHS	MHS	(1)	(2)	(1)	(2)
IU	None	1	0	0	0·11	0·41	0	0·06	0·37	0	5·78	14·4	2·26	
IR	No FS	0	0·56	0	0·10	0·41	0				7·54	14·4	1·73	
IIU	None	0·33	0·45	0·08	0·05	0·46	0·03	0·02	0·39	0·02	6·77	14·4	1·92	
IIR	No FS	0	0·63	0·06	0·05	0·45	0·03				7·60	14·4	1·72	
IIIU	None	0	0·68	0·16	0·03	0·48	0·05	0	0·40	0·03	7·41	14·4	1·76	1·61
IIIR	No PHS	0	0	1	0·02	0·23	0·07				8·66	14·4	1·51	1·38

† These values are independent of restrictions on male selection.
‡ (1) Derived from APPENDIX 2. (2) Derived from $N_e = 4N_mN_f/(N_m + N_f)$ (Wright, 1931).
§ With N_e derived from APPENDIX 2 and (1) $L = 3·83$; (2) $L = 4·18$.

TABLE 6

Estimated q(.,.:.) *and effective population size* (N_e) *for a variety of mating systems in juvenile MOET schemes*

Mating system	Restrictions on selection	q(m.m:.)			q(m.f:.)			q(f.f:.)†			N_e‡		Inbreeding rate§ (% per annum)	
		FS	PHS	MHS	FS	PHS	MHS	FS	PHS	MHS	(1)	(2)	(1)	(2)
IU	None	1	0	0	0·11	0·36	0	0·09	0·33	0	5·89	14·4	4·24	
IR	No FS	0	0·45	0	0·11	0·35	0				7·99	14·4	3·13	
IIU	None	0·33	0·36	0·14	0·06	0·41	0·04	0·03	0·36	0·03	6·94	14·4	3·60	
IIR	No FS	0	0·51	0·09	0·06	0·40	0·04				8·01	14·4	3·12	
IIIU	None	0	0·56	0·24	0·03	0·44	0·07	0	0·38	0·04	7·61	14·4	3·29	2·82
IIIR	No PHS	0	0	1	0·03	0·22	0·08				8·53	14·4	2·93	2·51

† These values are independent of restrictions on male selection.
‡ (1) Derived from APPENDIX 2. (2) Derived from $N_e = 4N_mN_f/(N_m + N_f)$ (Wright, 1931).
§ With N_e derived from APPENDIX 2 and (1) $L = 2$; (2) $L = 2·33$.

Rates of inbreeding

The frequencies of sib co-selection for the basic systems are given in Table 5 for adult MOET nucleus schemes and Table 6 for juvenile MOET nucleus schemes. To assess the frequencies it is helpful to consider $2q(..:F) + q(..;M) + q(..:P)$ for (m.m), (m.f) and (f.f) sampling, since this is a measure of the new inbreeding coming into the system each generation from each pathway.

For both adult and juvenile MOET nucleus schemes, the new inbreeding decreases when moving from system IU, through system IIU to system IIIU irrespective of whether (m.m), (m.f) or (f.f) sampling is considered (even though in (f.f) sampling q(f.f:U) remains near constant). When restricted, the new inbreeding for system IIR is less than system IR for (m.f) sampling but the reverse is true for (m.m) sampling. When the effective population sizes are calculated using these probabilities and the theory developed in APPENDIX 2, the effective population sizes (N_e) rank IU < IIU < IIIU. With restriction on male full-sib selection the difference between systems IR and IIR diminishes but the ranking is unchanged.

System IIIR has been included for

comparison although its restriction involves paternal half-sibs rather than full-sibs. When this restriction is made because of the structure of the scheme, males chosen are all maternal half-sibs. Even though the new inbreeding is thus greater on the (m.m) pathway the great reduction in new inbreeding on the (m.f) pathway more than compensates, so that N_e is increased under the restriction.

An important result is the 1·5- to 2·5-fold discrepancy between the estimate of N_e from the theory in APPENDIX 2 and estimates derived from the formula of Wright (1931).

The estimates of N_e presented for the schemes considered, give inbreeding rates of 0·065 per generation for the adult MOET nucleus scheme, using systems IR and IIR, and 0·063 for the juvenile MOET scheme operating the same mating systems (slightly faster for schemes using system IIIU). Thus, per generation, juvenile schemes apparently have the smaller inbreeding rate; however, because of the major difference in generation interval, per annum, the inbreeding rate is about twice as great in the juvenile schemes as in the adult schemes.

If, as with rates of progress, the rates of inbreeding in system IIR and IIIU are considered at an equal selection proportion of males as system IR, the inbreeding benefits of systems IIR and IIIU are emphasized. The changes in estimates of q(m.m:P) and q(m.m:M) and the resulting N_e are shown in Figures 3 and 4 for systems IIR and IIIU, respectively. At the same selection proportion in adult MOET nucleus schemes N_e = 7·56, 7·93 and 8·12 for systems IR, IIR and IIIU, respectively, and in juvenile MOET nucleus schemes the corresponding figures are N_e = 7·98, 8·41 and 8·58.

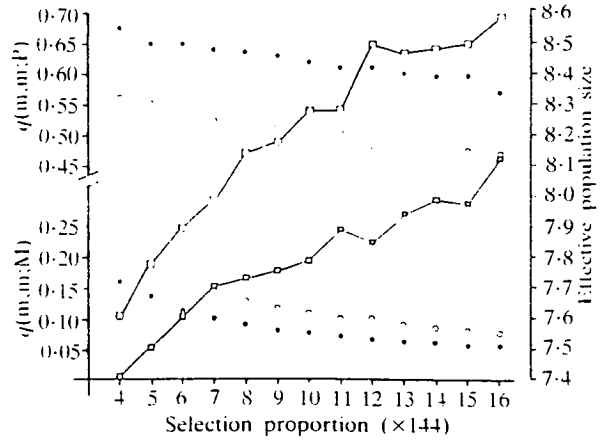


FIG. 3. Effective population size, (⊟ and □), q(m.m:M) and q(m.m:P) (● and ○) according to selection proportion for system IIIU in adult (closed symbols) and juvenile (open symbols) MOET nucleus schemes.

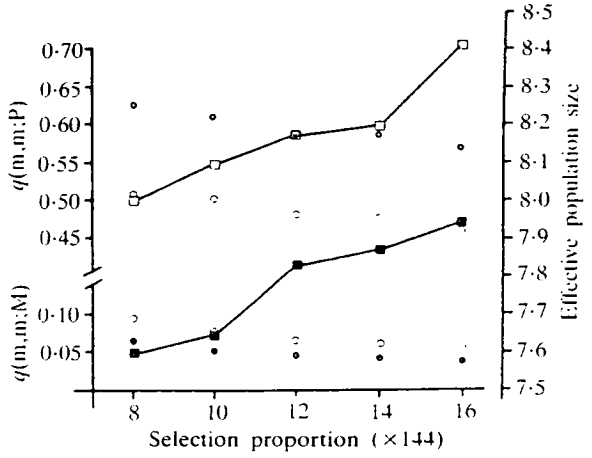


FIG. 4. Effective population size, (■ and □), q(m.m:M) and q(m.m:P) (● and ○) according to selection proportion for system IIR in adult (closed symbols) and juvenile (open symbols) MOET nucleus schemes.

Variation in h^2

The effect of varying h^2 on the accuracy of selection in adult MOET nucleus schemes using mating systems I and III is shown in Table 7. Accuracy increased with h^2 and the small advantage of system I over system III gradually diminished: when h^2 = 1, the accuracies of both systems are 1 for females and close to, but less than, 0·707 for males (the breeding value of the male's dam would be known but some error in the breeding value of its sire would remain). The correlations among paternal relatives decreased with h^2. In contrast, the maternal half-sib correlations of system III, increased with h^2 for males but increased and then decreased for females. When h^2 = 1, in both mating systems, these equal the coefficient of kinship for females and are close to twice the coefficient for males.

In accord with the negligible effect of heritability on the ranking of the mating systems for accuracy, the heritability made little difference to the relative rates of genetic progress achievable by the various mating systems (see Table 8).

The effect of h^2 on N_e is shown in Table 9. As h^2 increased so did N_e. The change in N_e was most evident for system IR: for h^2 = 0·1, IR had the lowest estimate of N_e (compared with IIR and IIIU) whereas for h^2 = 0·7, IR had the greatest estimate. The advantage of IIR over IIIU increased with h^2. When compared at equal selection proportions the systems always ranked, in ascending order of N_e, IR, IIR and IIIU, although the differences between them tended to diminish as h^2 increased.

TABLE 7

Effect of h² *on the accuracy and correlations among relatives for mating systems I and III*

	Heritability			
	0·1	0·25	0·5	0·7
System I				
♀ accuracy	0·504	0·657	0·787	0·868
♂ accuracy	0·443	0·557	0·626	0·654
♀,♀ FS correlation	0·896	0·835	0·731	0·643
♀,♀ PHS correlation	0·585	0·483	0·385	0·327
♂,♂ PHS correlation	0·697	0·623	0·569	0·545
System III				
♀ accuracy	0·490	0·652	0·784	0·867
♂ accuracy	0·426	0·548	0·619	0·649
♀,♀ PHS correlation	0·561	0·473	0·381	0·326
♀,♀ MHS correlation	0·330	0·363	0·348	0·317
♂,♂ PHS correlation	0·671	0·611	0·559	0·538
♂,♂ MHS correlation	0·329	0·388	0·441	0·462

DISCUSSION

The results presented here have (i) established the importance of considering the mating system in MOET nucleus schemes, (ii) quantified rates of inbreeding for such systems, (iii) removed some of the biases in rates of response (with finite resources) contained in previously published work (Nicholas and Smith, 1983; Woolliams and Smith, 1988) and (iv) shown some effects of heritability on the rates of response and inbreeding.

The mating system operated in a MOET nucleus scheme can affect the rates of genetic progress obtained. This was established by considering various changes to the mating system proposed by Nicholas and Smith (1983), with the aim of reducing the number of full-sib progeny in favour of creating maternal half-sibs. For each distinct full-sib pair removed, an additional paternal half-sib pair and a maternal half-sib pair were introduced, i.e. a net gain in relationships amongst offspring but with no coefficient of kinships greater than ¼ (at least in generation 1). Thus, the amount of information on each dam and each sire as measured by the number of their offspring remained constant.

The changes considered did not improve progress by increasing the accuracy of breeding value estimation. Indeed, Table 1 shows that the accuracy for individuals is

TABLE 8

Rates of genetic progress (△G) in the example adult MOET nucleus schemes according to h² *for different mating systems*

System	*L*	Selection proportion for males	△*G*(× 1000) (phenotypic s.d. per annum)			
			h^2 = 0·1	h^2 = 0·25	h^2 = 0·5	h^2 = 0·7
IR	3·83	0·111	49	102	169	217
IIR	3·83	0·056	52	111	184	236
IIR (relaxed)†	3·83	0·111	49	102	171	220
IIIU	3·83	0·028	55	117	195	250
IIIU (relaxed)†	3·83	0·111	50	104	176	225

† Proportion of males selected (from among which breeding males are randomly selected) is increased to equal that for IR.

TABLE 9

Effective population size (N_e) *of adult MOET nucleus schemes according to* h^2 *for different mating systems*

Heritability (h^2)	Mating system				
	IR	IIR	IIR (relaxed)†	IIIU	IIIU (relaxed)†
0·1	6·98	7·16	7·40	7·07	7·63
0·25	7·54	7·60	7·94	7·41	8·12
0·50	8·15	8·11	8·46	7·84	8·63
0·7	8·55	8·44	8·82	8·09	8·96

† Proportion of males selected (from among which breeding males are randomly selected) is increased to equal that for IR.

progressively reduced from system I to II, to III. An explanation for this comes by noting that the systems considered are analogous to balanced incomplete block designs with dams as blocks and sires as treatments (systems I and III representing degenerate forms, split-plot and complete factorial designs respectively). Over this class of designs with the numbers of dams, sires and offspring constant, it can be shown (J. A. Woolliams, unpublished results) that the accuracy of evaluating the mean breeding value of the offspring is conserved (i.e. $E(\hat{G}_i \Sigma_j \hat{G}_j)$ = constant where $\hat{G}_1, \hat{G}_2 \ldots \hat{G}_n$ are the estimated breeding values for the n offspring). This indicates the collective sense in which information is conserved by the designs. Since in systems II and III families are not discrete, information on one offspring gives information not only on its sire and dam, but through them, on all other sires and dams and hence on all other offspring. Thus, two unrelated individuals have a covariance (albeit small, with correlations of magnitude $< 0{\cdot}01$) between their estimates of breeding value. Consequently, there is a redistribution of information resulting in net reductions in the accuracy of estimating a single individual's breeding value. In system III, for example, this covariance is always positive mainly due to the influence of individuals who are simultaneously maternal half-sibs of one and paternal half-sibs of the other.

Small gains in selection intensity in schemes of finite size were obtained even using the same selection proportions in each system. Tables 1 and 2 show that the average pairwise index correlation among the offspring remained almost constant by the changes in mating system. It appears, therefore, that increased selection intensity may be achievable for the same average correlation by having more uniform weaker relationships amongst the population rather than small groups of stronger relationships. This is found to be so, for example, in the extreme case where all individuals are uniformly correlated with one another compared with independent families with uniform intra-family correlation (considered by Hill (1976) and Rawlings (1976)) when the average pairwise correlation is equal.

The second, and most significant, gain in progress arose from comparing the different mating systems at similar rates of inbreeding. Since at the same selection proportion, systems II and III had lower estimated rates of inbreeding than system I, the proportion selected could be decreased (intensity increased) in these systems to obtain comparable rates of inbreeding.

The inbreeding rate in nucleus schemes is thus a major consideration in their effectiveness, particularly for MOET nucleus schemes where inbreeding may critically affect the production of good-quality embryos by the females. The need for a method of estimating inbreeding under selection led to the approach developed in APPENDIX 2. The method can be viewed as a generalization of Wright (1931) and although difficulties lie in the estimation of the $q(.,:.)$ the subsequent calculation of the largest eigenvalue is a straightforward computational problem. Nevertheless, it must be recognized that the method presented is likely to overestimate the true value of N_e (N. Wray, personal communication). Over a period of time assumptions 2 and 3 presented will be broken with the latter likely to prove the more serious. Proliferation of offspring from the more superior of the selected individuals (passing on their relative advantage to their offspring) will further narrow the genetic base and so reduce N_e. The method presented must, therefore, be viewed as a first-order approximation which will still lead to

overestimation. N. Wray (personal communication) has developed second-order methods including grand-parent contributions in an attempt to overcome this problem.

The study has shown that applying the formula of Wright (1931) in MOET selection schemes is inappropriate leading to overestimation of N_e and a failure to discriminate among mating systems and selection procedures. This should not be surprising since the assumptions under which the formula was derived were that family sizes prior to the choice of breeding stock have a Poisson distribution and that the choice is then made at random. In selection this latter assumption would require no correlation of the index among relatives, a condition very far from met when family indices are used. Even though the schemes considered in this study have assumed equal family sizes prior to selection (thereby reducing the potential for inbreeding) the formula of Wright (1931) overestimates N_e roughly two-fold (in other circumstances the overestimate can be as great as three-fold; J. Woolliams, unpublished results).

In the absence of selection with random choice of breeding stock, the different mating systems (IU, IIU and IIIU) have no effect on N_e. This can be shown either for equal or for Poisson family size by noting matrix A (see APPENDIX 2) remains unchanged by the different mating systems. Colleau (1985) considers IIU with Poisson family sizes and arrives at a formula differing from that of Wright (1931); this is due to an error in his derivation and when corrected the formula returns to that of Wright. Therefore, it can be concluded that it is the imposition of selection, introducing correlations of index values among relatives, that creates the differences in N_e among the mating systems. In contrast, the restrictions on full-sib selection will affect N_e with or without selection (Wright's formula assumes that no restrictions exist).

Burrows (1984a and 1984b) attempted to take selection into account in the calculation of inbreeding rates but his approximations developed from Hill (1976) and Rawlings (1976), are unsatisfactory when applied to the schemes presented here. The formulae of Burrows asymptotically approach his definition of the rate of inbreeding as (i) the numbers of offspring before and after selection tend to infinity (for any value of the index correlations among relatives) and (ii) the correlations among relatives tend to zero. However, the formulae are inadequate for two reasons. First, some of the estimated frequencies of co-selection of relatives derived from his formulae are gross overestimates: this is most evident for q(m.m:P) in adult schemes with mating systems IR, IIR and IIIU where probabilities of 0·9, 1·4 and 2·2 respectively are derived (cf. Table 5). Secondly, the equating of inbreeding rate with the average pairwise coancestry of selected individuals put forward by Burrows is inappropriate when males and females are selected differently. From Table 5 it can be seen that if only (m.f) pairings are considered (which estimates the average inbreeding of the first generation of offspring) the effect of avoiding full-sib males is hardly accounted for. If the unweighted average of all (m.m) (m.f) and (f.f) pairings are considered, N_e is consistently overestimated (to be expected since whilst (f.f) pairings are the most frequent, individual females contribute far fewer genes to the next generation than individual males). If equal weightings are given to the average pairwise coancestry of (m.m) pairs, (m.f) pairs and (f.f) pairs calculated separately then the resulting estimate of N_e is less than those presented. Some of the problems described can be overcome (J. A. Woolliams, unpublished results) but will not be discussed further in this paper.

A further observation of note is the slightly lower estimate of inbreeding rate per generation obtained in juvenile schemes (in all but IIIR) compared with adult schemes. This is despite the fact that the total reliance on pedigree information in the former scheme implies that if any candidate (male or female) is selected then so are all its female full-sibs. However, this property, tending to increase inbreeding, is more than offset by a greater emphasis placed on dam, rather than sire, information in juvenile schemes (due to the greater accuracy of females in adult schemes, see METHODS). This reduces the degree to

which paternal half-sib males are selected, an important factor in determining the rate of inbreeding.

The results from varying the heritability suggested that when considered at similar rates of inbreeding, the factorial mating system (III) was likely to be most beneficial for low heritability. Nevertheless, for all values of h^2, system III appeared advantageous. Variation in t has not been considered but it is unlikely greatly to alter the results since it only affects the pooling of information from the two lactations of the dams of candidates.

Much of the previously published work (Nicholas and Smith, 1983; Woolliams and Smith, 1988; Woolliams, 1989) on MOET nucleus schemes have used selection intensities that if accounting for finite size do not account for the correlation of index values. The results of this paper show that allowing for the combination of finite size herds and the correlation of index values reduces expected progress to approximately 0·85 of that for large schemes. This is still greater than the figure of 0·65 derived by Juga and Mäki-Tanila (1987) from simulation studies. This discrepancy could be due to the reduction in the genetic variance from inbreeding or the effect of linkage disequilibrium (Bulmer, 1971) or stochastic variables. Wray and Hill (1989) using simulation to compare one generation responses (similar to that used here) and asymptotic rates of response found good concordance in the ranking of breeding schemes. However, these authors only considered nested mating systems. It is important, therefore that the conclusions arrived at in this paper are simulated to establish long-term response.

In practice, the changes in mating system can be implemented immediately. The results presented only consider discrete generations and would require confirmation with overlapping generations which are likely to be encountered commercially. It remains to be seen whether complete factorial mating (system III) can be achieved without significantly increasing the generation interval. However, systems such as IIR do not sustain any such penalty in the practical operation of a MOET nucleus breeding scheme. Such systems offer increased rates of genetic progress whilst simultaneously decreasing rates of inbreeding.

ACKNOWLEDGEMENTS

I am very grateful to Dr R. Thompson and Dr B. McGuirk for helpful discussions and comments and to a referee for suggesting some evaluation of the effect of the heritability on the results presented.

REFERENCES

BECKER, W. A. 1975. *Manual of Quantitative Genetics.* 3rd ed. Washington State University, Pullman, Washington.

BULMER, M. G. 1971. The effect of selection on genetic variability. *American Naturalist* **105:** 201-211.

BURROWS, P. M. 1984a. Inbreeding under selection from unrelated families. *Biometrics* **40:** 357-366.

BURROWS, P. M. 1984b. Inbreeding under selection from related families. *Biometrics* **40:** 895-906.

COLLEAU, J. J. 1985. [Genetic improvement by embryo transfer within selection nuclei in dairy cattle.] *Génétique Sélection Évolution* **17:** 499-537.

GOWE, R. S., ROBERTSON, A. and LATTER, B. D. H. 1959. Environment and poultry breeding problems. 5. The design of poultry control strains. *Poultry Science* **38:** 462-471.

HILL, W. G. 1976. Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32:** 889-902.

JUGA, J. and MÄKI-TANILA, A. 1987. Genetic change in nucleus breeding dairy herd using embryo transfer. *Acta Agriculturae Scandinavica* **37:** 511-519.

LAND, R. B. and HILL, W. G. 1975. The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Animal Production* **21:** 1-12.

NICHOLAS, F. W. and SMITH, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

RAWLINGS, J. O. 1976. Order statistics for a special class of unequally correlated multinormal variates. *Biometrics* **32:** 875-887.

WOOLLIAMS, J. A. 1989. The value of cloning in MOET nucleus breeding schemes for dairy cattle. *Animal Production* **48:** 31-35.

WOOLLIAMS, J. A. and SMITH, C. 1988. The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46:** 333-345.

WRAY, N. R. and HILL, W. G. 1989. Asymptotic rates of response from index selection. *Animal Production* **49:** In press.

WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics* **16:** 97-159.

(*Received* 18 *August* 1988—*Accepted* 29 *January* 1989)

APPENDIX 1

(with R. Thompson, AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS)

Calculations of variances and covariances of estimated breeding values in MOET schemes

Consider estimation of breeding value from the information on sires, dams and offspring. From theoretical results of BLUP the prediction error variance/covariance matrix for the estimated breeding values (V) of all individuals considered is given by $(R^{-1} + G^{-1})^{-1}$ where G is the variance/covariance matrix of the true breeding values for all individuals and R is the residual variance/covariance matrix of individuals with observations (i.e. so that $R + G$ for individuals with observations is the phenotypic variance/covariance matrix).

Consider the partition of $R^{-1} + G^{-1}$ into segments corresponding to sires, dams and offspring (S, D and O components respectively with dimensions s, d and o).

$$X = R^{-1} + G^{-1} = \begin{pmatrix} X_{SS} & X_{SD} & X_{SO} \\ X'_{SD} & X_{DD} & X_{DO} \\ X'_{SO} & X'_{DO} & X_{OO} \end{pmatrix}$$

It is required to calculate the segment of V corresponding (in position) to X_{OO} (i.e. V_{OO}). First premultiply by T where

$$T = \begin{pmatrix} I_s & O & -X_{SO}X_{OO}^{-1} \\ O & I_d & -X_{DO}X_{OO}^{-1} \\ O & O & X_{OO}^{-1} \end{pmatrix}$$

where I_n is the unit matrix of dimension n.

$$\text{with result } Y = \begin{pmatrix} Y_{SS} & Y_{SD} & O \\ Y'_{SD} & Y_{DD} & O \\ X_{OO}^{-1}X'_{SO} & X_{OO}^{-1}X'_{DO} & I_O \end{pmatrix}$$

where $Y_{SS} = X_{SS} - X_{SO}X_{OO}^{-1}X'_{SO}$
$Y_{SD} = X_{SD} - X_{SO}X_{OO}^{-1}X'_{DO}$
$Y_{DD} = X_{DD} - X_{DO}X_{OO}^{-1}X'_{DO}$

The transformed matrix is easier to invert and it can then be shown (proof by multiplication) that

$$V_{OO} = X_{OO}^{-1} + X_{OO}^{-1}(X'_{SO}\ X'_{DO}) \begin{pmatrix} Z_{SS} & Z_{SD} \\ Z'_{SD} & Z_{DD} \end{pmatrix} \begin{pmatrix} X_{SO} \\ X_{DO} \end{pmatrix} X_{OO}^{-1} \quad (1)$$

where

$$\begin{pmatrix} Z_{SS} & Z_{SD} \\ Z'_{SD} & Z_{DD} \end{pmatrix} = \begin{pmatrix} Y_{SS} & Y_{SD} \\ Y'_{SD} & Y_{DD} \end{pmatrix}^{-1}$$

By multiplication it can be shown that

$$Z_{SS} = (Y_{SS} - Y_{SD}\ Y_{DD}^{-1}\ Y'_{SD})^{-1}$$
$$Z_{SD} = -Z_{SS}\ Y_{SD}\ Y_{DD}^{-1}$$
$$Z_{DD} = Y_{DD}^{-1} - Y_{DD}^{-1}\ Y'_{SD}\ Z_{SS}\ Y_{SD}\ Y_{DD}^{-1}.$$

Thus multiplying (1) out

$$V_{OO} = X_{OO}^{-1} + X_{OO}^{-1}\ X'_{DO}\ Y_{DD}^{-1}\ X_{DO}\ X_{OO}^{-1} + X_{OO}^{-1}(X'_{SO} - X'_{DO}\ Y_{DD}^{-1}\ Y'_{SD}) Z_{SS}\ (X_{SO} - Y_{SD}\ Y_{DD}^{-1}\ X_{DO})\ X_{OO}^{-1}.$$

V_{OO} can be viewed as the sum of three components: (a) the first as the prediction error for the offspring conditional on the breeding value of dam and sire; (b) the second as the prediction error of the offspring arising from prediction errors in the dams breeding value conditional on the sires breeding value; and (c) the third as the prediction error of the offspring arising from prediction errors in the sires breeding value.

Consider now the information available in an adult MOET scheme. For simplicity assume that only female offspring are produced and each dam has n offspring in full-sib families of size f. Let each sire be mated to m dams. Thus total number of offspring per sire = mf. Each element of X is either the sum of elements of R^{-1} and G^{-1} or an element of G^{-1} only (e.g. for males where no direct observation is available).

Therefore

$X_{SS} = (2 + mf)(2h^2)^{-1}I_s$: contribution from G^{-1} only
$X_{DD} = [(\frac{1}{2}(1 + t) - h^2)^{-1} + (2 + n)(2h^2)^{-1}]\ i_d$: contribution from both R^{-1} and G^{-1}
$X_{OO} = [(1 - h^2)^{-1} + (h^2/2)^{-1}]\ I_o$: contribution from both R^{-1} and G^{-1}
$X_{DO} = (h^2)^{-1}\ N_{DO}$: contribution from G^{-1} only
$X_{SO} = -(h^2)^{-1}\ N_{SO}$: contribution from G^{-1} only
$X_{SD} = (2h^2)^{-1}\ N_{SD}$: contribution from G^{-1} only

where N_{DO}, N_{SO} and N_{SD} are offspring incidence matrices of dimensions $d \times o$, $s \times o$ and $s \times d$ respectively, and

$(N_{DO})_{ij} = 1$ if offspring j was from dam i; 0 otherwise
$(N_{SO})_{ij} = 1$ if offspring j was from sire i; 0 otherwise
$(N_{SD})_{ij} = f$ if sire i and dam j were mated; 0 otherwise.

Proof of these formulae is again by multiplication.

Thus,

$$Y_{SS} = (4 - 2h^2 + h^2\ md)[2h^2(2 - h^2)]^{-1}I_s$$
$$Y_{DD} = [2(1 + t)(2 - h^2) + nh^2\ (1 + t - 2h^2)]\ [2h^2(1 + t - 2h^2)(2 - h^2)]^{-1}I_d$$
$$Y_{SD} = [2(2 - h^2)]^{-1}N_{SD}$$

From these simple matrices all others can be derived and the problem of matrix inversion has been reduced to the inversion of the $s \times s$ matrix $Y_{SS} - Y_{SD}\ Y_{DD}^{-1}\ Y'_{SD}$ to obtain Z_{SS}. Once this has been achieved the prediction error variances can be derived for any subset of individuals. The variances and covariances of estimated breeding values can then be calculated by subtracting the prediction error variance/covariance from the corresponding genetic variance/covariance.

The formulae generalize in a straightforward manner to include male offspring.

APPENDIX 2

Assessment of inbreeding rates

Inbreeding rates have been calculated under the following assumptions.

1. Mating among selected individuals apart from restrictions imposed by the mating plans (i.e. fixed number of offspring per sire and per dam and fixed number of dams/sire and sires/dam) is at random.
2. Genetic parameters remain constant.
3. The selection process at time t is independent of the selection processes on all preceding occasions.

Assumption 2 will not be valid in the long term. Perhaps more seriously, assumption 3 is clearly invalid

for selection procedures since superior stock will tend to have superior offspring. However, the objective is to improve on current estimates of inbreeding rates with family indices (e.g. Nicholas and Smith, 1983) and it will be seen that even with these assumptions substantial improvements in estimating inbreeding rates can be made.

Each generation of selection introduces new potential for inbreeding by selection of full-sibs and half-sibs (maternal and paternal). If amongst all selected females and males there were no such relationships there would be no inbreeding i.e. no individuals with genes at a given locus that are identical by descent. Thus, the rate of inbreeding depends on the frequency with which full-sibs and half-sibs are chosen.

Let $p(\text{m.f:}t)$ be the probability that in generation t a gene taken from a selected male (m) and a gene taken from a selected female (f) are identical by descent, where the male and female are chosen at random from among the selected population. Similarly, let the probabilities of identity by descent when the sampling is from two distinct randomly chosen selected individuals of the same sex be $p(\text{m.m:}t)$ and $p(\text{f.f:}t)$. Furthermore, let $q(\text{m.f:}x)$ be the probabilities that the sampled male and female be full-sibs (x = F), paternal half-sibs (x = P), maternal half-sibs (x = M) or unrelated (x = U), with $q(\text{m.m:}x)$ and $q(\text{f.f:}x)$ similarly defined. From assumption 2 these are independent of t.

The inbreeding of an individual sampled at random in generation $t+1$, from assumption 1, is equal to $p(\text{m.f:}t)$.

From assumption 3

$$\begin{aligned}p(\text{m.f:}t) = {} & q(\text{m.f:F})(\tfrac{1}{2}p(\text{m.f:}t-1) + \tfrac{1}{4}p(\text{m.f:}t-2) + \tfrac{1}{4}) \\ & + q(\text{m.f:P})(\tfrac{1}{2}p(\text{m.f:}t-1) + \tfrac{1}{4}p(\text{f.f:}t-1) \\ & \quad + \tfrac{1}{8}p(\text{m.f:}t-2) + \tfrac{1}{8}) \\ & + q(\text{m.f:M})(\tfrac{1}{2}p(\text{m.f:}t-1) + \tfrac{1}{4}p(\text{m.m:}t-1) \\ & \quad + \tfrac{1}{8}p(\text{m.f:}t-2) + \tfrac{1}{8}) \\ & + q(\text{m.f:U})(\tfrac{1}{2}p(\text{m.f:}t-1) + \tfrac{1}{4}p(\text{m.m:}t-1) \\ & \quad + \tfrac{1}{4}p(\text{f.f:}t-1))\end{aligned}$$

$p(\text{m.m:}t)$ and $p(\text{f.f:}t)$ are derived similarly with $q(\text{m.f:}x)$ substituted by $q(\text{m.m:}x)$ and $q(\text{f.f:}x)$ respectively. Assumption 3 is required since the parents of selected individuals in generation t are assumed to be randomly sampled from the selected individuals of generation $t-1$.

Thus let $h_t = (1 - p(\text{m.m:}t),\ 1 - p(\text{m.f:}t),\ 1 - p(\text{f.f:}t),\ 1 - p(\text{m.f:}t-1))'$ then $h_t = Ah_{t-1}$ where the 4×4 matrix A has elements (scanned row by row):

$$\begin{gathered}\tfrac{1}{4}(q(\text{m.m:M}) + q(\text{m.m:U})),\ \tfrac{1}{2},\ \tfrac{1}{4}(q(\text{m.m:P}) + q(\text{m.m:U})), \\ \tfrac{1}{8}(2q(\text{m.m:F}) + q(\text{m.m:P}) + q(\text{m.m:M})), \\ \tfrac{1}{4}(q(\text{m.f:M}) + q(\text{m.f:U})),\ \tfrac{1}{2},\ \tfrac{1}{4}(q(\text{m.f:P}) + q(\text{m.f:U})), \\ \tfrac{1}{8}(2q(\text{m.f:F}) + q(\text{m.f:P}) + q(\text{m.f:M})), \\ \tfrac{1}{4}(q(\text{f.f:M}) + q(\text{f.f:U})),\ \tfrac{1}{2},\ \tfrac{1}{4}(q(\text{f.f:P}) + q(\text{f.f:U})), \\ \tfrac{1}{8}(2q(\text{f.f:F}) + q(\text{f.f:P}) + q(\text{f.f:M})),\ 0,\ 1,\ 0 \text{ and } 0.\end{gathered}$$

Thus in general $h_t = A^t h_0$.

By calculating the eigenvalues and eigenvectors of A the asymptotic rate of change of h can be found. From theory for a matrix of the form of A a real eigenvalue exists, the largest (λ_{max}) being less than or equal to 1 and greater or equal to $\frac{1}{4}(1 + \sqrt{5})$. The estimated effective population size (N_e) equals $\frac{1}{2}(1 - \lambda\max)^{-1}$.

As an example of the method consider N_m males mated at random to N_f females with N_f/N_m females/male, each mating producing one male and one female offspring. Selection of female parents for the next generation is automatic and one male parent is chosen at random from each sire group. For large N_m and N_f, ignoring terms small compared to N_m^{-1} and N_f^{-1}: $q(\text{m.m:U}) = 1$, $q(\text{m.f:F}) = N_f^{-1}$, $q(\text{m.f:P}) = N_m^{-1} - N_f^{-1}$, $q(\text{m.f:U}) = 1 - N_m^{-1}$, $q(\text{f.f:P}) = N_m^{-1} - N_f^{-1}$, $q(\text{f.f:U}) = 1 - N_m^{-1} + N_f^{-1}$ and all other q probabilities are 0.

Thus A =

$$\begin{pmatrix} \tfrac{1}{4} & \tfrac{1}{2} & \tfrac{1}{4} & 0 \\ \tfrac{1}{4}(1 - N_m^{-1}) & \tfrac{1}{2} & \tfrac{1}{4}(1 - N_f^{-1}) & \tfrac{1}{8}(N_m^{-1} + N_f^{-1}) \\ \tfrac{1}{4}(1 - N_m^{-1} + N_f^{-1}) & \tfrac{1}{2} & \tfrac{1}{4} & \tfrac{1}{8}(N_m^{-1} - N_f^{-1}) \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Including only first order terms in N_m^{-1} and N_f^{-1}, λ_{max} is a root of the equation

$$\lambda^4 - \lambda^3 + \tfrac{1}{16}(N_m^{-1} - N_f^{-1})\lambda^2 + \tfrac{1}{32}(N_m^{-1} + 3N_f^{-1})\lambda.$$

with solution $\lambda_{max} = 1 - (\frac{3}{32})N_m^{-1} - (\frac{1}{32})N_f^{-1}$

giving $$\frac{1}{2N_e} = \frac{3}{32}N_m^{-1} + \frac{1}{32}N_f^{-1}$$

the result derived by Gowe, Robertson and Latter (1959).

Paper 15

Anim. Prod. 1989, 48: 31-35
© 1989 British Society of Animal Production

0003-3561/89/8260003131$02·00

THE VALUE OF CLONING IN MOET NUCLEUS BREEDING SCHEMES FOR DAIRY CATTLE

J. A. WOOLLIAMS

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

ABSTRACT

The value of cloning in MOET nucleus breeding schemes has to be considered within the constraint of fixed resources. Under this constraint cloning was found to add to genetic progress only when (i) the heritability is low and (ii) it is used at the expense of a reduction in the number of bull families. This course would exacerbate inbreeding and other potential problems with MOET. All other options for using clones lead to a reduction in genetic progress due to a loss of selection intensity that is not made up for by gains in selection accuracy.

INTRODUCTION

THE technique for embryo splitting to give identical individuals was first reported in ruminants by Willadsen (1979). Nicholas and Smith (1983) were quick to consider the technique for the benefits it could achieve in breeding schemes particularly those involving multiple ovulation and embryo transfer (MOET). They viewed the technique as being beneficial; they recognized the fact that replication leads to greater accuracy in estimating an individual's merit and its use, they argued, would thus speed progress for the same rate of inbreeding — the cost of this gain being more transfers of embryos.

This paper reviews this finding under the more realistic constraint that the number of transfers, in determining the numbers of recipients and candidates, is likely to be the limiting factor in MOET breeding schemes. It assumes, as does Nicholas and Smith (1983), that the keeping of more than a single member of a clone family because of the major reductions in genetic variance and increases in inbreeding this would cause, would not be commercially sound. Under these circumstances the production of two-clone families for all genetically distinct individuals, for example, can only be done at the expense of halving the number of such individuals, thus reducing the selection intensity; further cloning further erodes intensity. It will be shown that under the constraint of fixed resources the usefulness of embryo splitting is at best equivocal.

METHODS

The essential details of the MOET nucleus scheme considered are selection for a sex-limited trait (T) in which selection occurs after one observation on the trait for females, and the information used in selection also includes two further observations from the female parent. Thus, with T as milk yield, the narrow heritability (the proportion of the total variance of T that is additive genetic) h^2 = 0·25 and a generation interval L = 3·67 years is equivalent to the adult scheme given by Nicholas and Smith (1983). To generalize, the results have been considered for a range of h^2 values: namely 0·05, 0·1, 0·25 and 0·5 and 0·7. The genetic variation among clones is likely to be greater than h^2 and the broad heritability (H^2) has been taken to be H^2 = 1·2 h^2. The repeatability (r) of T used to obtain the variance of repeat measurements from the parent, has been assumed r = 1·4 h^2. For h^2 = 0·25 these are consistent with the values assumed by Nicholas and Smith (1983). For simplicity it has been assumed that the difference between h^2 and H^2 is caused by dominance rather than

epistatic interactions. Selection practice has been assumed consistent with Nicholas and Smith (1983) in that females will be selected on an index as individuals with no restriction on the numbers selected per full-sibship whereas males will be selected on an index for full-sibship merit and subsequently selected within sibships at random with one male per sibship. It has been assumed that family groups produced by MOET will be balanced between males and females, although the potential of embryo sexing (a non-random departure from this assumption) is considered briefly later.

The resources of the scheme have been set at 256 transferred offspring produced per selection round and selection intensities (i (n_1, n_2); the selection intensity obtained by selecting the best n_1 from n_2 individuals which are normally distributed of known mean and unit variance) reflect this. Values used in the paper are from the tables of Becker (1975).

Any MOET nucleus scheme for using these resources depends on four parameters: s, the number of bulls used per round; d, the mating ratio, the number of donors mated to each bull; f, the total number of full-sibs, male and female produced per mating; and c, the number of clones per offspring. These obey the constraint $sdfc = 256$. Since 256 is a power of 2, each parameter is a power of 2.

Let r_F and r_M be the correlation of the indices with breeding values for T for female and male selection respectively; these are functions of d, f and c and their derivation is given in APPENDIX 1. The rate of additive genetic progress per annum measured in genetic standard deviations ($\triangle G$) is given by

$$\triangle G = 0{\cdot}5\ [i(sd,\tfrac{1}{2}\mathrm{sdf})r_F(d,f,c) + i(s,sd)r_M(d,f,c)]/L.$$

To assess the effect of changing c at the expense of either changes in d, f or s, $\triangle G$ was calculated for

s = 1, 2 and 4
d = 2, 4, 8 and 16
f = 4 up to $256/(sd)$
and $c = 256/(sdf)$.

In the results changes in c have been considered with compensating changes in either of s, d or f while keeping the other two constant. A base class of (s, d, f, c) = (2, 8, 8, 2) has been used for comparison for Figure 1. This scheme does not give the greatest value of $\triangle G$ for a particular value of c but compromises between $\triangle G$ and the size of genetic base, i.e. number of genetically distinct individuals selected.

RESULTS AND DISCUSSION

Effect of changing s

Changes in s are different from those in d and f since the selection proportions for both males and females remain constant, though small changes in i_F and i_M occur through changes in numbers available for selection. Additional clones increase both r_F and r_M (Table 1). Under these circumstances the gain in accuracy obtained from additional clones is beneficial to $\triangle G$ when h^2 is low. This benefit decreases as h^2 increases, so that when $h^2 = 0{\cdot}7$ the gain in accuracy is insufficient to make up the loss of intensity from selecting in a reduced population of distinct genotypes (see Figure 1a).

The effect of reducing the number of bull families will, however, lead to increased rates of inbreeding and loss of genetic variation, thus exacerbating perceived problems with MOET (Woolliams and Smith, 1988). When values of d and f are held constant but the resources whilst remaining fixed are scaled up to accommodate large values of s, the effect of complimentary changes in s and c are similar to those from small s. r_F and r_M remain as calculated in this paper but i_F and i_M tend to their large sample values. Results are shown (for $h^2 = 0{\cdot}25$) in Table 1. In these circumstances with large numbers of bull families, it may be considered acceptable when selecting for traits of low or medium heritability, to reduce the number of families in order to obtain benefits from cloning.

Effect of changing d

Decreases in d to compensate for increases in c reduce the proportion of males selected but not the proportion of females selected and the population size for selection is reduced for both sexes. From Figure 1b and Table 1 it is seen that any gains in accuracy

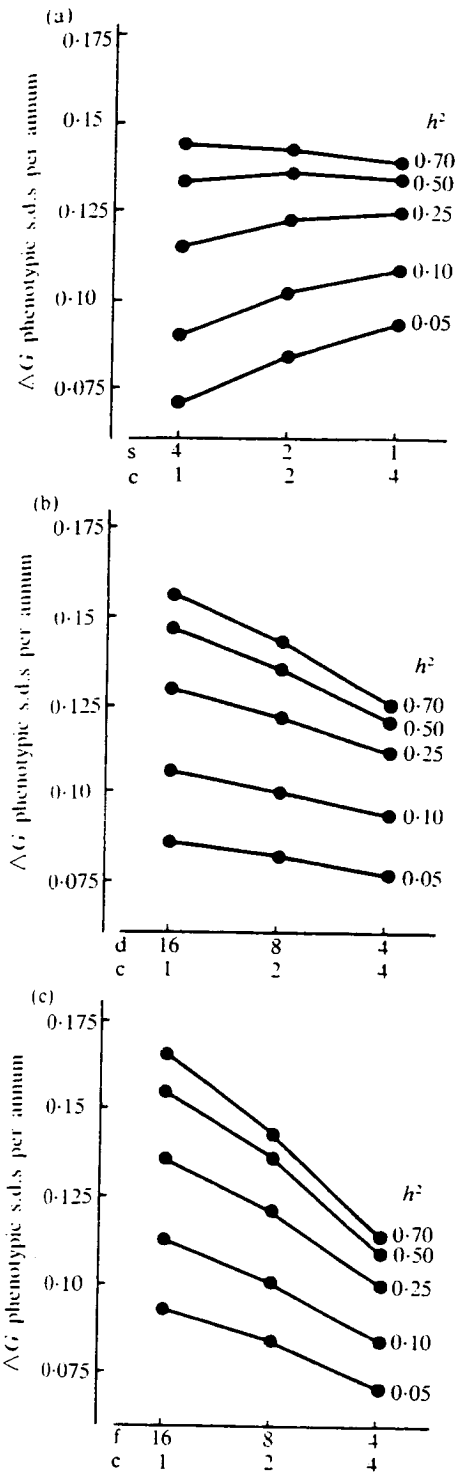


FIG. 1. Annual genetic progress expected in an adult MOET scheme of fixed resources: (a) $d = f = 8$; (b) $s = 2$, $f = 8$; (c) $s = 2$, $d = 8$.

resulting from the clones do not offset losses in selection intensity. For females the additional clones provide direct estimates of genetic merit on the candidates themselves and so increase r_F. In contrast, for males, r_M is decreased by additional clones for some values of s and f if $h^2 \geq 0{\cdot}10$. This occurs because the gain in accuracy in already existing sibs does not offset the loss of information from distinct half-sibs.

Effect on changing f

Increases in c at the expense of f reduce the proportion of females selected but does not change the proportion or the number of males selected. Whilst again r_F is increased by cloning, r_M is reduced since the loss of distinct full-sibs is not compensated for by additional information on existing full-sibs (Table 1). The net result as shown in Figure 1c is a drop in ΔG as more clones are produced at the expense of full-sibs.

General

It may at first sight appear that embryo sexing could overcome some of the disadvantages since only female embryos need be split and only a single male embryo need be produced per full-sib family. However, embryo sexing in this context is a method of freeing resources for use and the question returns to that addressed by this paper — with information coming from a fixed number of females what is the optimum division among half-sibs, full-sibs and clones? This paper suggests that clones should make a very limited contribution.

One of the concerns over the estimates of rates of genetic progress obtainable by MOET nucleus breeding schemes is the degree to which correlations of index values among half-sib and full-sib relatives erode the selection intensities assumed in the calculations (Hill, 1976; Rawlings, 1976). Cloning, by giving more precise information on individual candidates, reduces this correlation among females and so reduces the problem (Table 1). Among male full-sibships the correlation is also generally, but not always, decreased.

The expected reductions in selection intensity are not calculable from current

TABLE 1

The values for i_F, i_M, r_F, r_M *and* ΔG *for different parameters* (s, d, f, c) *together with the correlation among index values of half-sibs and full-sibs for both the male and female indices for* $h^2 = 0{\cdot}25$

Scheme	i_F	r_F	i_M	r_M	ΔG† (1) phenotypic s.d. per annum	ΔG† (2)‡ phenotypic s.d. per annum	Correlation among index values§ Female full-sibs	Female half-sibs	Male half-sibs
(*s*, *d*, *f*, *c*)									
(2, 8, 8, 2)	1·252	0·720	1·525	0·577	0·122	0·127	0·765	0·443	0·582
(1, 8, 8, 4)	1·235	0·784	1·424	0·598	0·124	0·135	0·698	0·391	0·544
(4, 8, 8, 1)	1·264	0·650	1·583	0·547	0·115	0·118	0·828	0·493	0·607
(2, 4, 8, 4)	1·235	0·778	1·138	0·583	0·111	0·118	0·694	0·386	0·506
(2, 16, 8, 1)	1·264	0·664	1·858	0·570	0·130	0·134	0·836	0·509	0·635
(2, 8, 4, 4)	0·779	0·774	1·525	0·562	0·100	0·105	0·691	0·385	0·598
(2, 8, 16, 1)	1·632	0·673	1·525	0·586	0·136	0·142	0·840	0·504	0·568

† (1) assumes the correct finite selection intensities (i_F and i_M) whilst (2) assumes large sample values.
‡ For large n: $i(n,2n) = 0{\cdot}798$, $i(n,4n) = 1{\cdot}271$, $i(n,8n) = 1{\cdot}647$, $i(n,16n) = 1{\cdot}967$.
§ Values depend on d, f and c (APPENDIX 2) and assume large s.

theory since Hill (1976) and Rawlings (1976) only consider one level of grouping not the two levels (half- and full-sibs) found in MOET female selection. However, the information given by these authors makes it possible to give systematic over- and under-estimates of potential progress: (i) the effect of the half-sib grouping among male full-sibships can be calculated directly: (ii) if the correlation among full-sib females is assumed to be 1 and the selection intensity is calculated appropriate for the half-sib correlation among females, then, when combined with (i) expected progress is underestimated: (iii) if the correlation among half-sibs is assumed to be 0 and the selection intensity is calculated appropriate for the full-sib correlation among females, then, when combined with (i) expected progress is overestimated. Using the methods shown in APPENDIX 2 these calculations have been made for the examples given in Table 1, and the results (not shown) indicate that the correlations have no qualitative effect on the conclusions for changing s and f. For changes in d, whilst the rankings of the over- and under-estimates of potential progress remain as before the ranges of possible progress do not exclude an intermediate optimum (i.e. $c = 2$). Nevertheless, resources freed by reducing d to produce more clones may be better used, should technology allow, in increasing f.

The modelling has assumed a constant multiplicative relationship between additive genetic and total genetic variance ($H^2 = 1{\cdot}2\ h^2$). The value of cloning will increase as H^2 tends to h^2 since the clone family means will more closely estimate breeding value but the general conclusions remain valid.

In conclusion the results show that in the context of MOET nucleus schemes embryo splitting cannot be assumed to be beneficial to genetic progress. In the small scheme considered and in selection for milk yield, cloning is advantageous only when it is used at the cost of reducing bull families, a cost that would exacerbate many of the perceived problems with MOET.

REFERENCES

BECKER, W. A. 1975. *Manual of Quantitative Genetics* 3rd ed. Washington State University, Pullman, Washington.

HILL, W. G. 1976. Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32**: 889-902.

NICHOLAS, F. W. and SMITH, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36**: 341-353.

RAWLINGS, J. O. 1976. Order statistics for a special class of unequally correlated multinormal variates. *Biometrics* **32**: 875-887.

WILLADSEN, S. M. 1979. A method for culture of micromanipulated sheep embryos and its use to produce monozygotic twins. *Nature, London* **277**: 298-300.

WOOLLIAMS, J. A. and SMITH, C. 1988. The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46**: 333-345.

(*Received* 19 *February* 1988—*Accepted* 5 *July* 1988)

APPENDIX 1

Calculation of r_M *and* r_F

Let σ_A^2, σ_D^2, σ_P^2 be the additive genetic, dominance and phenotypic variances for T. Thus $\sigma_A^2/\sigma_P^2 = h^2$ and $(\sigma_A^2 + \sigma_D^2)/\sigma_P^2 = 1{\cdot}2h^2$ (see assumptions in METHODS). The index weights are given by $b = V^{-1}\mathbf{g}$ where V is the variance-covariance matrix of the information going into the index and $\mathbf{g}$ is the vector of its covariances with breeding value. Thus, when scaled by σ_P^2 for females

$V =$	Dam	Self	Full-sibs	Half-sibs
	v_{11}	v_{12}	v_{13}	0
	v_{12}	v_{22}	v_{23}	v_{24}
	v_{13}	v_{23}	v_{33}	v_{34}
	0	v_{24}	v_{34}	v_{44}

where $v_{11} = 0{\cdot}5\ (1 + 1{\cdot}4h^2)$
$v_{12} = v_{13} = 0{\cdot}5h^2$
$v_{22} = 1{\cdot}2h^2 + (1 - 1{\cdot}2h^2)/c$
$v_{23} = 0{\cdot}55h^2$
$v_{24} = v_{34} = 0{\cdot}25h^2$
$v_{33} = 0{\cdot}55h^2 + (v_{22} - 0{\cdot}55h^2)/(0{\cdot}5f - 1))$
$v_{44} = 0{\cdot}25h^2 + (0{\cdot}3dh^2 + (v_{22} - 0{\cdot}55h^2)/(0{\cdot}5fd)$

(note that for v_{11} repeatability of yield has been assumed to be $1{\cdot}4h^2$ and that in v_{23}, v_{33} and v_{44} the covariance of full sibs is $0{\cdot}5\sigma_A^2 + 0{\cdot}25\sigma_D^2 = 0{\cdot}55h^2\sigma_P^2$)

and $\mathbf{g} = 0{\cdot}5h^2,\ h^2,\ 0{\cdot}5h^2,\ 0{\cdot}25h^2$

The correlation of the selection index with the breeding value for T is given by

$$r_I = b'\mathbf{g}\ (b'Vbh^2)^{-1/2} = (\mathbf{g}'V^{-1}\ \mathbf{g}/h^2)^{1/2}$$

For males r_M is obtained using the same equation as that given for r_F but after deleting the second row and column of V and the second column of $\mathbf{g}$, and substituting $0{\cdot}5f$ for $0{\cdot}5f - 1$ in v_{33}.

APPENDIX 2

Correlation of index values among relatives

These are calculated assuming infinite population sizes produced by increasing s, but maintaining the values of d, f and c.

For females the covariance matrix between two full-sib females is given by C_{FS} and between two half-sibs is given by C_{HS}

C_{FS} =

		Full-sib 1			
		Dam	Self	Full-sib mean	Half-sib mean
Full-sib 2	Dam	v_{11}	v_{12}	v_{13}	0
	Self	v_{12}	a_{22}	a_{23}	v_{24}
	Full-sib mean	v_{13}	a_{23}	a_{33}	v_{34}
	Half-sib mean	0	v_{24}	v_{34}	v_{44}

where v_{ij} are equal to the elements of V given in APPENDIX 1 and

$a_{22} = 0{\cdot}55h^2$
$a_{23} = [0{\cdot}55\ (0{\cdot}5f - 2)\ h^2 + v_{22}]/(0{\cdot}5f - 1)$
$a_{33} = [0{\cdot}55\ (0{\cdot}25f^2 - 1{\cdot}5f + 3)h^2$
$+ (0{\cdot}5f - 2)v_{22}]/(0{\cdot}5f - 1)^2$

(derived by noting sib 1 and sib 2 appear in the others full-sib mean and have identical relationships with their dam and half-sibs)

and C_{HS} =

		Half-sib 1			
		Dam	Self	Full-sib mean	Half-sib mean
Half-sib 2	Dam	0	0	0	x_{14}
	Self	0	x_{22}	x_{23}	x_{24}
	Full-sib mean	0	x_{23}	x_{33}	x_{34}
	Half-sib mean	x_{14}	x_{24}	x_{34}	x_{44}

where $x_{22} = x_{33} = x_{23} = 0{\cdot}25h^2$
$x_{14} = 0{\cdot}5h^2/(d - 1)$
$x_{24} = x_{34} = [v_{22} + 0{\cdot}55\ (0{\cdot}5f - 1)h^2$
$+0{\cdot}125\ (d - 2)fh^2]/[0{\cdot}5\ (d - 1)f]$
$x_{44} = [(d - 2)v_{22} + 0{\cdot}55\ (d - 2)\ (0{\cdot}5f - 1)h^2$
$+0{\cdot}125\ (d^2 - 3d + 3)fh^2]/[0{\cdot}5\ (d - 1)^2f]$

derived by noting that the dams are unrelated and that sib 1 and her full-sibs appear in the half-sib mean of sib 2 and *vice versa*. One consequence is a covariance between the dam of one sib and the half-sib mean of the other.

For males only C_{HS} is required.

C_{HS} =

		Half-sib 1		
		Dam	Full-sib mean	Half-sib mean
Half-sib 2	Dam	0	0	y_{13}
	Full-sib mean	0	y_{22}	y_{23}
	Half-sib mean	y_{13}	y_{23}	y_{33}

where $y_{22} = x_{33}$, $y_{33} = x_{44}$, $y_{23} = x_{34}$ and $y_{13} = x_{14}$

The index coefficients are given by $V^{-1}\mathbf{g}$ (APPENDIX 1) and the index has a variance of $\mathbf{g}V^{-1}V\ V^{-1}\mathbf{g} = \mathbf{g}V^{-1}\mathbf{g}$.

Correlation of selection index among relatives with covariance matrix C is then given by $(\mathbf{g}V^{-1}C\ V^{-1}\mathbf{g})/(\mathbf{g}V^{-1}\mathbf{g})$.

Paper 16

Anim. Prod. 1989, 48: 3-30

0003-3561/89/8395000350200

EMBRYO MANIPULATION IN CATTLE BREEDING AND PRODUCTION

J. A. WOOLLIAMS AND I. WILMUT
AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

ABSTRACT

Developments, both recent and potential, in procedures for manipulating embryos are described. The procedures considered include: embryo transfer, multiple ovulation and embryo recovery, recovery of oocytes, *in vitro* maturation (IVM) and fertilization (IVF) of oocytes, *in vitro* culture of zygotes, embryo splitting and nuclear transfer, embryo storage, embryo sexing, gene transfer and embryo stem cells. The impact of these procedures on breeding strategies such as multiple ovulation and embryo transfer (MOET) nucleus breeding schemes and progeny testing are discussed for both dairy and beef cattle.

For MOET nucleus schemes all these procedures have potential applications in producing maximal rates of genetic progress for a fixed rate of inbreeding. With the current effectiveness of the procedures, embryo sexing and nuclear transfer would have the most impact. The potential for increasing genetic progress through progeny testing is enhanced using multiple ovulation, embryo recovery and transfer in cows to breed bulls, but no other procedures appeared to offer major benefits. The efficiency of beef production from the dairy herd could be increased either by using IVM and IVF to produce more beef-type calves or, potentially, by cloning and embryo transfer, to produce pure beef calves. Procedures leading to the production of clone families would make an impact on the evaluation of genotypes and environments. Gene transfer may be used to modify the composition of milk including the production of pharmaceutical proteins, and to increase milk yield or the efficiency of lean meat production.

It is concluded that, although much further research is required, the procedures discussed will have major implications for the structure and organization of dairy and beef cattle herds over the next decade.

INTRODUCTION

JOHN HAMMOND'S interest in embryo transfer developed naturally from his pioneering work with Arthur Walton to establish artificial insemination (AI) of cattle in Great Britain. In the course of this work it was realized that the transfer of embryos from females that had been induced to have multiple ovulations could offer another means of obtaining genetic improvement and (at a time when specialist breeds were replacing dual-purpose breeds) increasing the number of beef calves (Hammond reviewed progress in 1950). In 1949 the Agricultural Research Council established the Unit of Animal Reproduction (UAR) under the direction of Hammond with a primary aim being to develop methods of embryo transfer. A unique contribution to the development of methods for embryo transfer and manipulation was made at the Unit and at the Unit of Reproductive Physiology and Biochemistry (URPB) (which the UAR became on Hammond's retirement). During the period from 1949 to 1986 (when URPB became part of the Institute of Animal Physiology and Genetics Research) reliable procedures for embryo transfer were first described (Rowson, Moor and Lawson, 1969), equipment was developed for non-surgical recovery (Newcomb, Christie and Rowson, 1978) and techniques were introduced for the freezing of embryos (Wilmut and Rowson, 1973), for the *in vitro* maturation of oocytes (Staigmiller and Moor, 1984), for embryo splitting and for nuclear transfer (Willadsen, 1982 and 1986). These, and other related procedures, may in time, make a contribution to cattle breeding and production comparable to that of AI.

This review will describe briefly each of the potential methods of cattle embryo manipulation, give an indication of when they will become available for commercial exploitation and consider their potential application. It will consider embryo sexing, but not semen sexing. The procedures to be considered and the relationship between them are summarized in Figure 1.

METHODS FOR ASSESSING GENETIC PROGRESS AND INBREEDING

Throughout this review both genetic progress and inbreeding have been considered per generation rather than per annum. This is because attention has been focussed on what the embryo procedures could offer within current breeding schemes (instead of a direct comparison between schemes). Part of the development of these techniques will involve their effective operation within as short a period of time as possible so as to maintain (or shorten) current generation intervals.

In the evaluation of new procedures in multiple ovulation and embryo transfer (MOET) nucleus herds, differing mating systems will be discussed which will be defined in terms of n_s, n_d, n_f, n_m and n_c denoting per generation, the numbers of sires used, dams used, female genotypes produced, male genotypes produced and the size of a clone family (e.g. $n_c = 1$ if neither cloning nor splitting is used). Comparisons among alternatives have been made with fixed resources (T) defined as the expected number of calves at the time of selection.

Genetic progress. The rates of genetic progress for milk yield that have been used in this review are either published estimates or derived using the formula

$$\Delta G = \tfrac{1}{2}(i_M r_M + i_F r_F)h$$

where ΔG denotes genetic progress in phenotypic standard deviations per generation: i denotes selection intensity; r denotes the correlation between the selection index and breeding value (termed selection accuracy): the subscripts M and F refer to males and females respectively; and h^2 is the narrow heritability of yield (the proportion of total variance that is additive genetic i.e. due to variation in breeding values). The values for r have been derived using methods for best linear unbiassed prediction (Woolliams, 1989b) and values for i have been calculated by simulation accounting for correlations of indices among relatives (Woolliams, 1989b). The calculations require the narrow heritability, broad heritability (H^2, the proportion of total variance that is due to variation in genotypes i.e. including non-additive genetic variation) and the repeatability of yield over two consecutive lactations and these have been assumed to be 0·25, 0·30 and 0·35 respectively. It is also assumed the genetic correlation of two consecutive lactations is 1.

Inbreeding. Inbreeding has been measured using the method of Woolliams (1989b) which accounts, in part, for the effect of selection on inbreeding rates. The measure increases towards the effective population size for random mating as the selection intensity tends to zero and is thus on a comparable scale with results using standard formulae (e.g. Wright, 1931; Gowe, Robertson and Latter, 1959). The greater the measure the smaller the rate of inbreeding is expected to be.

PRODUCTION AND TRANSFER OF EMBRYOS

Embryo transfer

Without the ability to transfer embryos from one cow to another none of the other procedures to be described would have any practical use. For this reason it is the lynch-pin of all other procedures.

There have been great changes in the methods of embryo transfer since the establishment of the first reliable procedures involving general anaesthetic and mid-ventral laparotomy (Rowson *et al.*, 1969). First, there was a change to the use of para-vertebral anaesthetic and a flank approach and more recently to non-surgical transfer (Newcomb, 1979). Experienced practitioners have been able to obtain pregnancies with around 66% of embryos using any of the methods of transfer. It is not obvious why procedures that initially failed to produce an acceptable proportion of pregnancies should now be found to be successful (early results reviewed by Newcomb *et al.*, 1978). There is

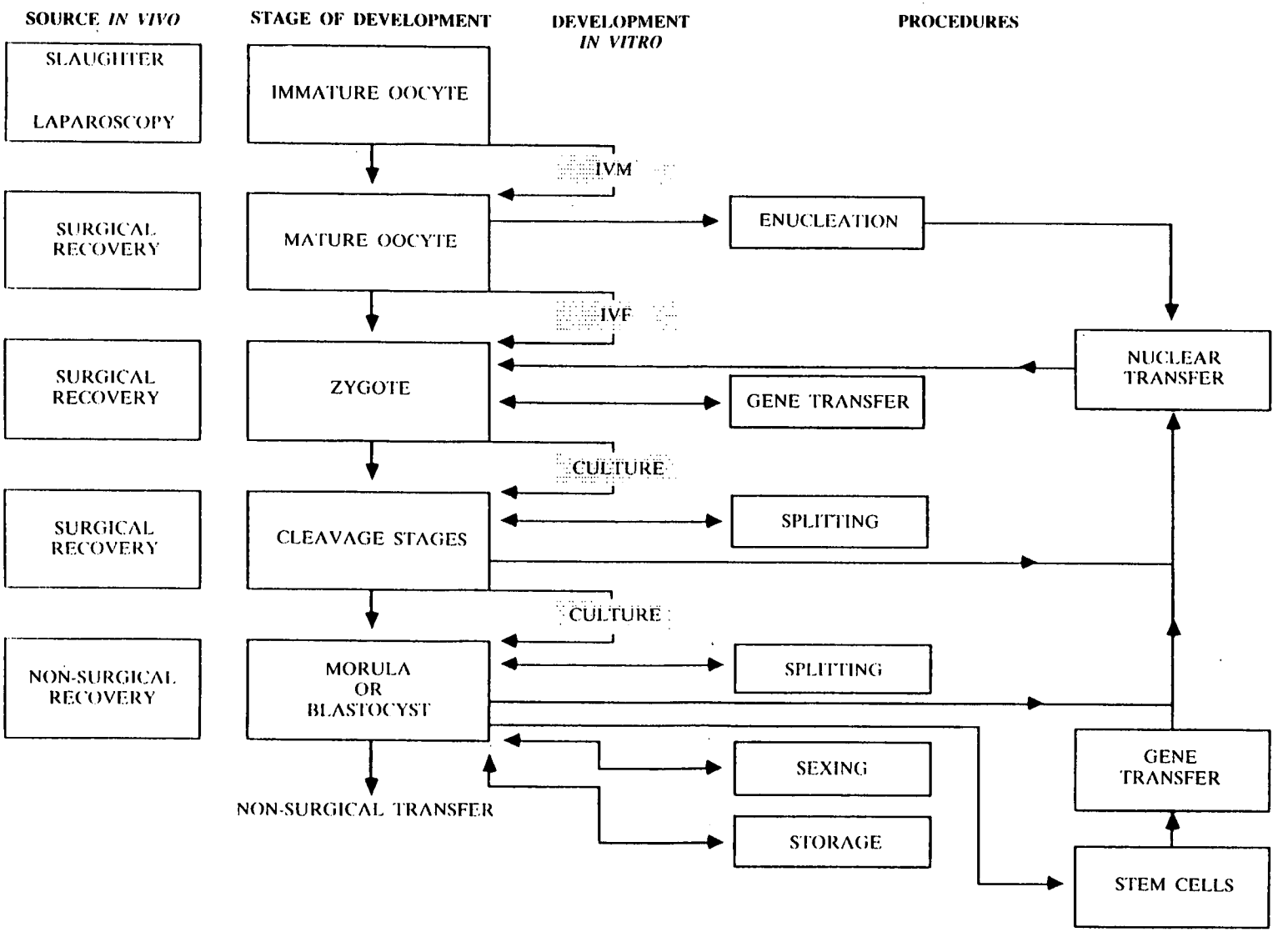


FIG. 1. Flow chart of the procedures for *in vitro* development and manipulation of oocytes and embryos prior to non-surgical transfer.

experimental evidence that conception rate increases with greater experience and practice (Rowe, Del Campo, Critser and Ginther, 1980). It may be very important to learn to handle the uterus with care.

In addition to experience, many factors influence the success of embryo transfer, by any of the established methods, including embryo stage and quality, the 'synchrony' between donor and recipient, and freezing and thawing (see later). Conception rate is also higher if the recipients are prepared by the transfer company (Christie, 1986). Given a good-quality embryo (graded 1 and 2 on a three-point scale by subjective assessment of morphology), that had not been frozen, more than 75% of recipients at a specialist centre would be expected to conceive, compared with only 67% of similar recipients on farms. This difference may reflect inexperience in carrying out critical parts of the procedures such as heat detection. It is very encouraging to note that, in general, the proportion of recipients calving is very similar to or higher than that achieved after a single AI (Hunter, 1984).

Multiple ovulation and embryo recovery

The development of reliable methods for non-surgical recovery of cattle embryos in the 1970s created entirely new opportunities in cattle breeding. Since then a very large number of cattle embryos have been recovered for transfer by commercial companies.

Success depends upon (i) controlling the time of oestrus and ovulation, (ii) inducing the animal to superovulate and (iii) recovering the embryos by passing sterile medium through the uterus.

Two approaches to the control of oestrus have been used: administration of progestagen to extend the luteal phase of the oestrus cycle, or administration of prostaglandin to induce premature regression of the animals own corpus luteum. There does not appear to be a major difference in their effectiveness (Gordon, 1982), but the use of prostaglandin has proved more popular.

Whichever method of oestrus control is used, gonadotrophin is given before the expected time of oestrus to promote follicle growth and so induce multiple ovulation. Gonadotrophin from two sources has been used very extensively: serum from pregnant mares (PMSG) and follicle stimulating hormone (FSH) prepared from horse or pig pituitaries. There is evidence that a single administration of PMSG continues to promote the growth of follicles even after the induction of oestrus, an effect attributed to the exceptionally long half-life of the hormone (Schams, Menzer, Schallenberger, Hoffmann, Hahn and Hahn, 1978). By contrast, FSH injections must be given either daily (equine) or twice daily (porcine). Despite this disadvantage, during the past decade there has been a gradual change from using PMSG to using FSH.

Non-surgical recovery depends upon passage along the uterine horn of a flexible catheter through which sterile medium can be passed to rinse out the embryos. This is achieved by passing a metal trochar and cannula through the cervix and along the uterus to the external bifurcation. The trochar is removed to allow a flexible catheter to be passed through the metal cannula and along the horn. Two catheters designed for non-surgical recovery have found commercial acceptance: two-way and three-way (see Newcomb, 1982). Both have a cuff that can be inflated to occlude the uterine lumen and a tube to allow the injection of fluid into the uterine lumen. In one case the fluid is recovered through the same tube, while in the second there is an additional tube for the recovery of medium.

The success of multiple ovulation and embryo recovery is very variable with yields of embryos from 0 to 40. This arises from wide variation in the induced ovulation rate (see Table 1 for example) and probably reflects variation in the number of follicles at the time of the gonadotrophin injection (Rajakoski, 1960). There has been a gradual increase in the number of embryos obtained from a donor at a single recovery. Typically a group of cows treated by an experienced laboratory would be expected to yield eight embryos per donor per recovery (Hasler, McCauley, Lathrop and Foote, 1987).

After multiple ovulation and embryo recovery the donors are allowed to return to

TABLE 1

Individual responses: number of eggs shed in a single cycle in 60 primiparous Friesian females all given 3120 i.u. pregnant mare's serum gonadotrophin from the same batch of the hormone preparation. The animals were given 0, 750, 1500, or 3000 i.u. human chorionic gonadotrophin, but this treatment was without effect on the number of eggs shed†

6	17	15	1	19	8	36	15	1	11
1	9	8	5	1	9	13	4	16	3
13	4	3	1	21	5	10	7	2	10
13	24	1	10	15	1	20	6	22	14
18	2	0	1	3	1	8	6	16	6
1	40	4	3	4	7	1	1	7	1

† Data from Mauleon *et al.* (1970).

oestrus before further treatment is initiated. As a result of this delay, multiple ovulation and embryo recovery can only be repeated every 6 weeks. There is a gradual reduction in the number of ovulations induced by successive treatments (Saumande, Chupin, Mariana, Ortavant and Mauleon, 1978). Use of different hormone preparations ameliorates this effect, but after a small number of treatments it is necesary to allow the animal to become pregnant.

There are particular difficulties in inducing multiple ovulation in older females and in such animals it may be more effective to attempt to recover the single embryo present after a natural oestrus. This may also be useful in order to maximize the yield of embryos from a particularly valuable animal.

An alternative source of embryos

Embryo recovery from a donor is limited by the ability of the female to produce embryos. In order to obtain access to the pool of oocytes (thousands) that remain in the ovary, attempts are being made to recover follicular oocytes, and to use culture methods for maturation, fertilization and embryonic development to a stage at which they can be transferred to a recipient. Oocytes are being obtained by laparoscopy from pre-ovulatory follicles and from slaughtered animals.

Recovery of oocytes by laparoscopy. Prior to slaughter the only way to obtain oocytes is through laparoscopic aspiration of follicles. This has been described in cattle and was used in an early trial to obtain 6·9 oocytes with normal morphology per donor per recovery (Lambert, Sirard, Bernard, Beland, Rioux, Leclerc, Menard and Bedoya, 1986). The yield would be expected to increase with experience.

Recovery is carried out after administration of gonadotrophin releasing hormone to induce maturation of the oocyte, but the oocyte must be cultured for 4 to 6 h to complete maturation. It seems possible that the yield could reach 10 oocytes per animal per recovery in superovulated animals. The feasibility of using this procedure on a practical scale has yet to be confirmed.

Oocyte maturation in vitro *(IVM).* In the interval between the onset of oestrus and ovulation, the oocyte undergoes changes within the ovary that are essential if it is to become capable of normal development. In recent years, methods have been established that support normal oocyte maturation in culture (Staigmiller and Moor, 1984; Lu, Gordon, Gallagher and McGovern, 1987).

Oocyte development is governed by the cumulus cells in which it is enclosed. These cells inhibit oocyte maturation throughout the life of the animal until the pre-ovulatory surge of LH in the cycle in which the follicle is to ovulate. Normal maturation in culture depends upon inducing the appropriate changes in cumulus cells. Oocytes enclosed within a cumulus mass are freed from translucent, vascular follicles 2 to 5 mm in diameter. They are then cultured in medium supplemented with serum and hormones for 24 to 26 h.

Two different approaches have been used to the development of effective media for oocyte maturation and both have limitations. One has employed oestrus cow serum, that is serum collected from cattle at the time of follicle maturation (Lu *et al.*, 1987). It is assumed that such serum contains all essential factors at an appropriate concentration. In the other, foetal calf serum has been included in the medium and pituitary gonadotrophins and steroids added at known concentrations

(Staigmiller and Moor, 1984). In this case, the assumption is that all of the important factors have been identified and are available. Surprisingly, the conclusion from a range of studies is that foetal calf serum can support the maturation of oocytes of other species, apparently as effectively as serum from oestrous females of the same species as the oocyte. An important research objective is to identify the essential factors and their optimal concentrations, but at present the empirical use of oestrus cow serum is proving valuable.

Both the current and future effectiveness of IVM is difficult to define, due in part to the paucity of published work. The great majority of oocytes complete meiosis in culture. However, even when transferred to inseminated recipients, not all have the ability to develop to blastocysts (Staigmiller and Moor, 1984). Future studies of IVM need to assess the ability of matured oocytes to develop to term when fertilized.

Fertilization in vitro (IVF). Before fertilization can occur *in vitro*, sperm have to undergo capacitation, a process which normally occurs in the female tract. Successful routines for IVF have been established by N. L. First and his colleagues at Madison. They have shown that it is important to control several factors. First, there is an optimum number of motile sperm with too great a number leading to penetration of the oocyte by more than one sperm. It is assumed that polyspermy is incompatible with normal development. Secondly, there are differences between bulls and even between ejaculates from the same bull in the response to treatment. Thirdly, sperm have a reduced life after freezing and thawing and the laboratory treatments must be completed within this time.

In practice a 'swim-up' procedure is used to increase the proportion of motile sperm (Parrish, Susko-Parrish, Leibfried-Rutledge, Critser, Eyestone and First, 1986) and heparin is used to induce capacitation (Parrish, Susko-Parrish, Winer and First, 1988). The 'swim-up' procedure involves pouring a layer of medium on top of thawed semen, and during a 1-h incubation only motile sperm are able to swim into the top layer of fluid. The top fluid is removed and exposed to heparin for 15 min. A known number of sperm are then added to oocytes in culture.

In many experiments the proportion of eggs fertilized has been assessed by fixing and staining the eggs for microscopic examination. Under ideal conditions over 90% of oocytes have been judged fertilized by morphological criteria, but there have been abnormalities of fertilization in up to 20% of these cases. It should not be assumed that the 70% of eggs that were judged to be fertilized normally had the potential to develop to term. Further published information on viability is required.

Embryo culture. Embryos must be at the morula or blastocyst stage before non-surgical transfer to a recipient. Development to these stages can be achieved either in a temporary recipient or in culture.

Rabbits and sheep have been used as temporary recipients for cattle embryos. While rabbits are able to support the development of four- and eight-cell embryos to a morula or a blastocyst, the development of one-cell embryos is limited (Lawson, Rowson and Adams, 1972). By contrast, a majority of one-cell embryos reach these stages in the oviduct of sheep (Willadsen, 1982). Almost 40% of the embryos are lost when transferred to a ewe, as they are not recovered from the recipient (Lu, Gordon, McGovern and Gallagher, 1988). The cost of using temporary recipients is not large as over 100 embryos can be transferred to each ewe. This method is in use.

The proportion of embryos developing in culture has been greatly increased by the use of co-culture systems. When cultured with epithelial cells from the cow oviduct, almost half the embryos reached the blastocyst stage (Eyestone, Vignieri and First, 1987). This technique may be capable of further improvement. The requirements of the embryo may change after early cleavage and it may be necessary to develop a two-stage method of culture, first using cells from the oviduct and then other cells, perhaps from the uterus. A system of culture would avoid the loss of embryos in the temporary recipient and would be more convenient.

The yield of embryos. By the combined use of IVM, IVF and culture in a sheep oviduct

approximately 20% of follicular oocytes became embryos suitable for transfer in one large trial (Lu *et al.*, 1988). The survival of these embryos after transfer is being determined in large-scale applications. The cost of an embryo produced in this way is less than that of embryos recovered from donors, perhaps £5 per embryo compared with £25 per embryo. In principle, there is a very large number of oocytes available for maturation as they can be obtained from females after slaughter.

There are now two practical procedures for the production of embryos: either non-surgical recovery from superovulated donors or culture. They have different potential applications.

PROCEDURES OF MANIPULATION

Embryo storage

Cattle embryos tolerate storage in a refrigerator or at ambient temperatures for several hours (Trounson, Willadsen, Rowson and Newcomb, 1976). This provides an opportunity to hold embryos for brief periods and transport them to other farms, but storage for longer periods requires freezing and thawing. The method for freezing depends upon the objective: to maximize survival or to simplify the procedure to allow transfer on farms without the use of a microscope.

To maximize survival after freezing and thawing, cooling and warming rate must be controlled, along with the choice and concentration of cryoprotective agent (see review by Wilmut, 1986). Conventional methods cool the embryos at around 0·3°C/min in the presence of 0·1 l glycerol per l of medium, to a temperature of a few degrees below the freezing point of the medium. The formation of ice crystals is induced by placing cold metal forceps against the outside of the container. At a temperature between −20 and −40°C the embryos are transferred to liquid nitrogen for storage at −196°C. Embryos cooled in this way should be warmed at several hundred °C/min, usually by shaking the container in warm water (Leibo, 1988). It is then essential to remove the glycerol before the embryo is transferred to a recipient. In the first procedures, this was achieved by serial dilution, but more recently the sugar sucrose has been used to reduce the quantity of glycerol in the embryo during dilution (Renard, Heyman and Ozil, 1982). Sucrose remains outside the cells of an embryo and so causes them to shrink. If the medium and embryo contain glycerol, such shrinkage reduces the quantity of glycerol, but not the concentration, in the embryo. Experience in the laboratory showed that passing the embryos from the medium containing glycerol to a medium containing sucrose allowed direct dilution of glycerol. This observation provided the basis for a method in which embryos are frozen in an insemination straw in the presence of glycerol. Another droplet of a medium that contains sucrose is placed in the same straw. After thawing, the straw is shaken in order to mix the drops of liquid and so make two changes in the environment of the egg: to reduce the concentration of glycerol and to introduce sucrose. This routine enables a technician trained in AI to thaw the embryo and, after allowing time for the removal of glycerol to occur in the straw, to transfer the embryo.

There has been no direct comparison of the two approaches, but it seems probable that survival of embryos is higher with the laboratory procedures. During extensive commercial use of the laboratory procedure, 74% of embryos survived freezing and thawing (W. B. Christie cited by Wilmut, 1986).

More recent attempts to simplify the procedures have used media containing higher concentrations of cryoprotective agents to prevent the formation of ice, an approach known as 'vitrification' (Rall and Fahy, 1985). In the reports published so far, the proportion of embryos surviving vitrification is lower than can be obtained with conventional procedures (Rall, Wood, Kirby and Whittingham, 1987)

Cloning

The ability to produce groups of identical individuals is of interest for both scientific and commercial reasons. This has lead to the

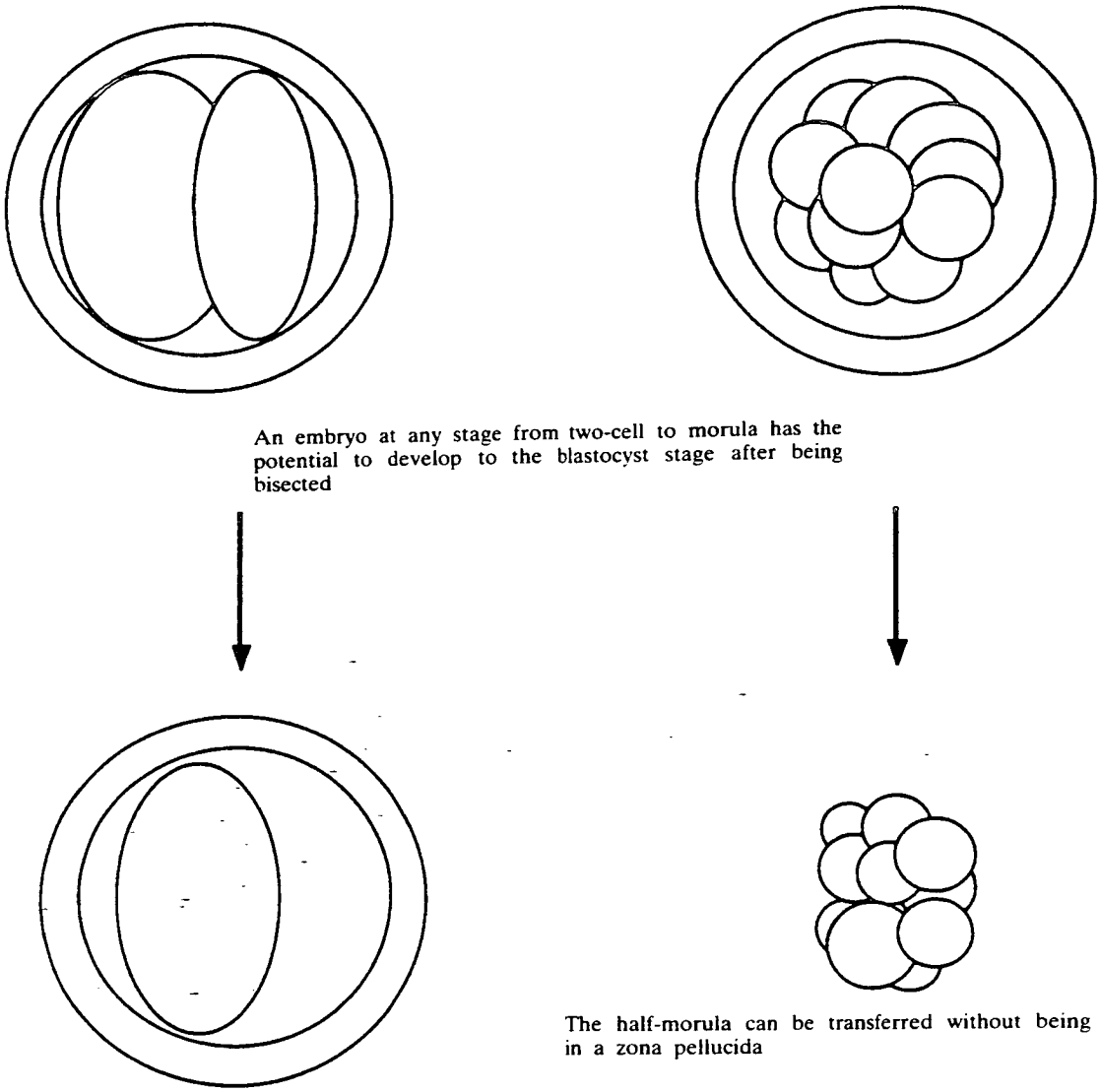


FIG. 2. Embryo splitting. Two embryos can be produced by splitting an embryo at any stage from two-cell to early blastocyst. After splitting of cleavage-stage embryos the halves must be placed in a zona pellucida for their protection. By contrast, morulae and blastocysts are able to survive transfer to a recipient without a zona. The additional zona is obtained from the ovaries of slaughtered animals.

development of two effective procedures: embryo splitting and nuclear transfer.

Embryo splitting (Figure 2). In a large proportion of cases, two embryos form if a cow embryo is bisected at any stage from two-cell up to, but not beyond, early blastocyst (see Willadsen, 1982). In a large commercial application of bisection procedures, 422 embryos were split and produced 441 pregnancies after transfer as single embryos,

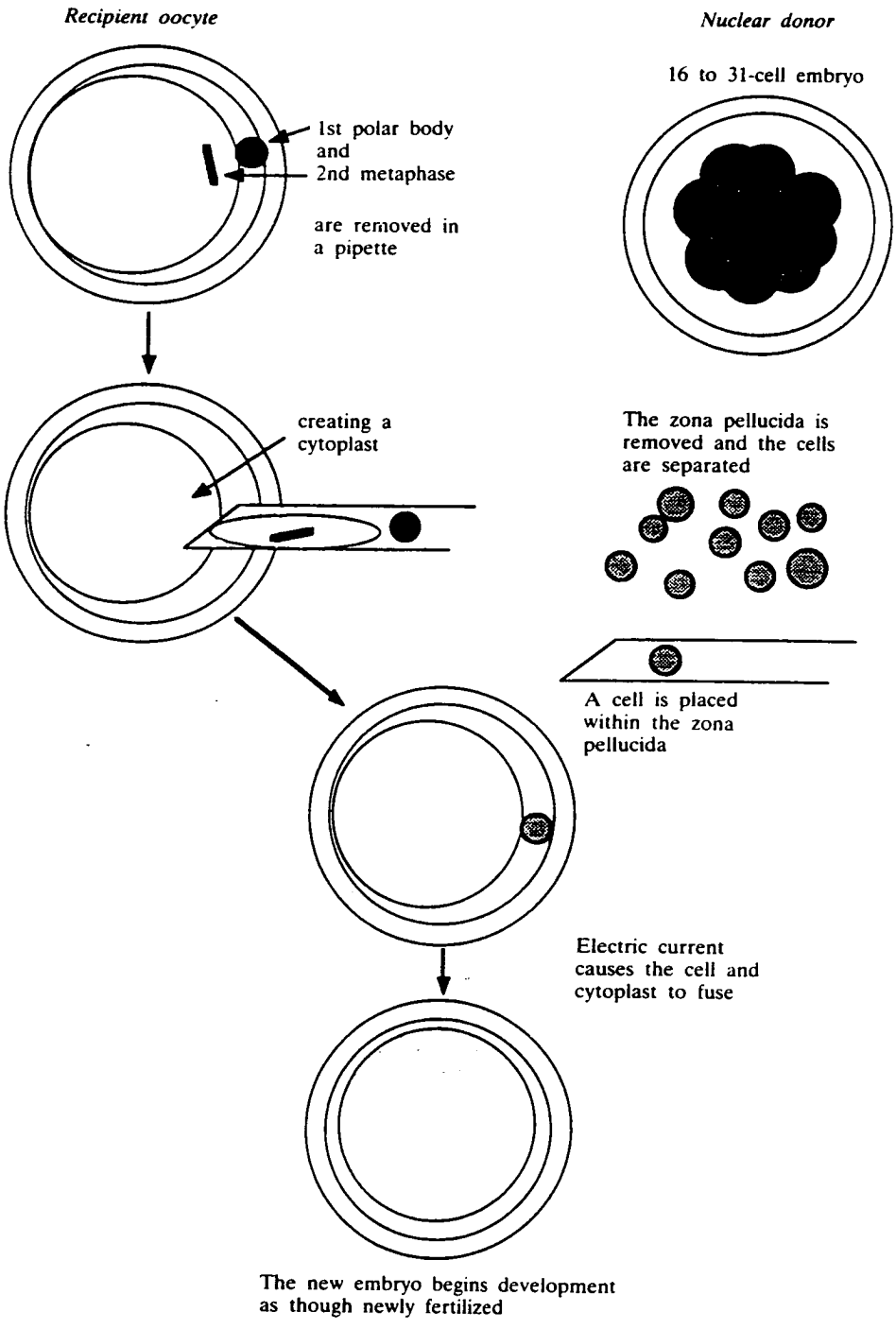


FIG. 3. Nuclear transfer. A nucleus is transferred by fusing a cell from an embryo to a secondary oocyte from which the chromosomes have been removed. The chromosomes are removed, by suction, into a pipette. The cell is placed within the zona pellucida by the same pipette, before electric current is passed to cause fusion of the two cells. The new embryo begins development as though newly fertilized.

whilst transfer of 515 intact embryos from the same donors produced 291 pregnancies (Leibo and Rall, 1987). Thus conception rate in the two groups of recipients was similar, being 52·4% after splitting compared with 56·5% with intact embryos. After splitting, the yield of calves was 1·05 times the number of embryos.

While this procedure has found commercial application there are three limitations. First, repeated splitting is not possible because the path of development is not affected by splitting. Both halves of the embryo form blastocysts at the time that the single blastocyst would have formed. Secondly, normal development does not occur if the embryo is split into more than four portions, apparently because each part of the embryo contains too few cells for normal development (Rossant, 1976). Thirdly, as splitting reduces the ability of the embryo to survive freezing and thawing, splitting and transfer must be carried out on the same day.

Nuclear transfer (Figure 3). These limitations of embryo splitting were overcome by the establishment of the technique for transferring nuclei from blastomeres to enucleated oocytes (Willadsen, 1986). Nuclear transfer resets the clock of development and the reconstituted embryo begins development as though it were a recently fertilized egg. The potential yield of embryos from each donor embryo has yet to be determined, but will almost certainly be more than four.

A recipient cytoplast is formed by removing the chromosomes from a recently ovulated oocyte. In the presence of drugs that breakdown the cytoskeleton of the cell, the first polar body and the neighbouring cytoplasm are drawn into a pipette without piercing the cell membrane. A nucleus is transferred by placing the donor cell against the cytoplast and passing electric currents to cause fusion of the cell and the cytoplast. Fusion is more likely to occur if the plane of the area of contact between the two bodies is perpendicular to that of the electric field. This alignment can be achieved by passing alternating current. The two bodies become polarized and roll to minimize resistance. A number of pulses of direct current are then passed to cause fusion.

There are several steps in nuclear transfer and losses occur at each step. In optimum conditions fusion can occur in about 90% of cases. The new embryos are embedded in agar and transferred to sheep for development to the blastocyst stage and of these only 60% are recovered from the ewe. There is little published information on the effect of stage of the donor embryo on the development of the embryo formed after nuclear transfer. In the first experiments nuclei were transferred from embryos with eight or 16 cells, but successful transfer has since been achieved with nuclei from embryos with 32 and 64 cells. It is anticipated that following transfer of nuclei from embryos at later stages of development fewer of the reconstituted embryos will be able to develop to blastocysts. At present, development to the morula or blastocyst stages occurs in around 40% of cases following transfer of nuclei in sheep (Smith and Wilmut, 1988). Therefore, if the donor embryo has 20 cells suitable for transfer, the yield would be around five embryos per donor embryo ($20 \times 0{\cdot}9 \times 0{\cdot}6 \times 0{\cdot}4$). In time, success rates will increase. If they were to reach 90% for each step and culture were to replace the temporary recipient, then the yield could reach 15 identical embryos per donor embryo. In addition it is probable that nuclear transfer could be repeated several times by using some of the new embryos as donors of nuclei. In this way it may be possible to generate clone families of several hundred or thousand calves from one donor embryo. These calculations provide only a naïve indication of the potential of the technique o nuclear transfer. It must not be assumed that embryos that develop to blastocyst are able to develop to term and it should be remembered that there have been very few reports on the ability of the embryos to develop to term.

It is a simplification to describe this procedure as nuclear transfer, because the whole of the cytoplasm of the donor cell is transferred. The effect of the cytoplasm on the development of the embryo or performance of the offspring is not known. The volume of a cell from the inner cell mass of a blastocyst is much less than that of the cytoplast, but it will contain organelles

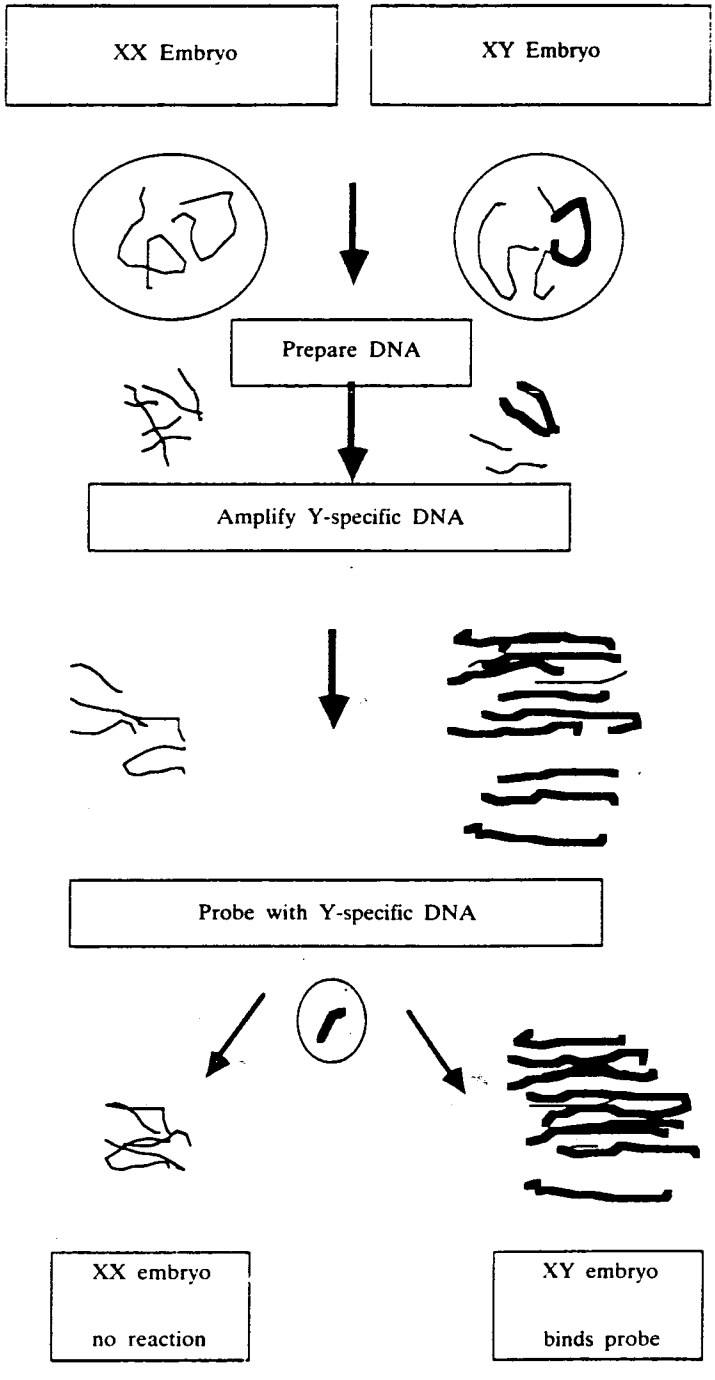


FIG. 4. Sexing. The sex of an embryo can be established by probing DNA from the embryo with DNA that is specific to the Y-chromosome. A number of cells are removed from the embryo by biopsy (not shown). The DNA is extracted from the cells and the specific sequences amplified by the polymerase chain reaction. The Y-specific probe is labelled with a biological label, such as biotin, and added to the preparation. The labelled DNA only binds to male embryos.

such as mitochondria. Mitochondrial DNA functions in a semi-independent manner and may effect metabolism. Heterogeneity in mitochondrial DNA has been reported (Freeman, 1988) and may explain discrepancies between progeny test results and maternal performance in dairy cattle (Freeman, 1988), but until methods of transferring nuclei without cytoplasm are developed a critical test of this suggestion is not possible. The importance of it is discussed in a later section.

In conclusion, embryo splitting is an established routine that requires only simple equipment, but it has a limited potential. By contrast, nuclear transfer requires much more research, but has a much greater potential.

Sexing of embryos (Figure 4)

Several approaches are being followed in the search for a method of identifying the sex of an embryo. These include karyotyping, using the H-Y antigen, measurement of sex-linked differences in enzyme activity and probing for DNA specific to the Y chromosome (reviewed by Wilmut and Smith, 1988). The last is the most promising method. Several laboratories have isolated DNA sequences that are specific to the Y chromosome and used them to make almost perfect predictions of the sex of the offspring (Jones, Singh and Edwards, 1987; Ellis, Bondioli, Williams, Pryor and Harpold, 1988). The commercial effectiveness of the procedure may depend upon how frequently the sequence is repeated on the Y chromosome.

Cells are removed from the embryo at the morula or blastocyst stage by micromanipulation. The loss of viability caused by the biopsy is likely to be small, but has still to be determined. DNA is extracted from the cells and sequences unique to the region in question are used as th template for amplification by the polymerase chain reaction. Several thousand copies of that particular part of the genome are synthesized. It is the development of this routine for rapid, specific amplification that permits identification of sex with only a small number of cells. The Y-specific probe is then used to identify the male embryos. The probe may be labelled with a biological label (e.g. biotin) or radioactive phosphorus.

This procedure could be rapid (within 2 h) and could be incorporated into any scheme

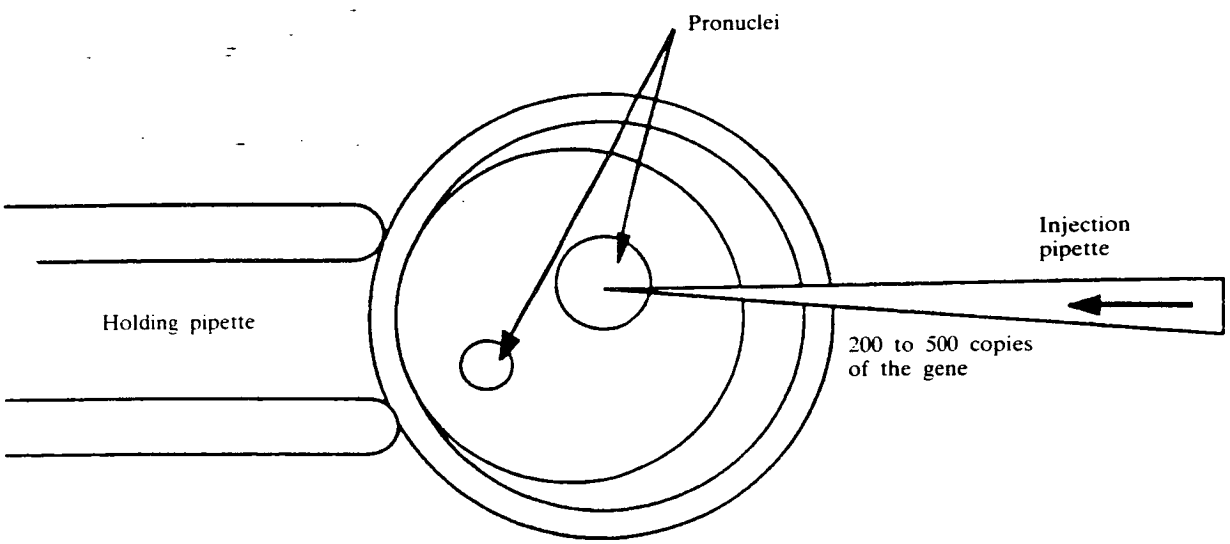


FIG. 5. Gene transfer. A gene is transferred by direct injection of 250 to 500 copies of the gene directly into a nucleus of an early embryo. The embryo shown is at the single cell stage and has two pronuclei: one derived from the oocyte, the other from the sperm. While the egg is held on the suction pipette the very fine injection pipette is inserted into a pronucleus and a small volume of fluid containing the DNA is injected.

involving embryo transfer, but it will add to the cost. Costs are incurred for the micromanipulation, laboratory consumables and for the waste of the embryos of the sex that is not required. If embryo splitting or nuclear transfer were being used to make clone families it would be desirable to know the sex of the donor embryo.

Gene transfer (Figure 5)

There are revolutionary techniques for the isolation of a gene from one animal, its multiplication and modification in the laboratory and its transfer to another animal, of the same, or a different species. These methods have been used extensively in mice (see Palmiter and Brinster, 1986) and modified for use in pigs, sheep and cattle (Hammer, Pursel, Rexroad, Wall, Bolt, Ebert, Palmiter and Brinster, 1985; Simons, Wilmut, Clark, Archibald, Bishop and Lathe, 1988; Biery, Bondioli and DeMayo, 1988). Three factors are important for the application of the procedure: efficient methods for incorporating genes into a host genome, the control of their expression after incorporation and, perhaps the most difficult, the identification of genes capable of having a desirable effect. The selection of suitable genes for transfer and methods for control of gene expression will be reviewed in a later section.

The most common procedure for gene incorporation in farm animals depends upon direct injection of a few hundred copies of a gene directly into a nucleus of an early embryo. The egg is held at the tip of a pipette by gentle suction while a very fine pipette is inserted into the nucleus. Distension of the nucleus confirms that buffer containing the DNA has been injected into the nucleus successfully. Usually the embryo has a single cell and the two pronuclei formed around the chromosomes of the sperm and egg are present. However, gene incorporation has been achieved by injection into a nucleus of two-cell embryos (Brinster, Chen, Trumbauer, Yagle and Palmiter, 1985; Pursel, Miller, Pinhert, Palmiter and Brinster, 1988). In the latter case the animal is almost certain to be a mosaic, with a mixture of cells only some of which carry the transgene. By contrast, following injection into pronuclei in mice, about 30% of lines were mosaic (Wilkie, Brinster and Palmiter, 1986).

In cattle, the pronuclei are obscured by cytoplasmic granules and it is necessary to centrifuge the embryos. Centrifugation at 10 to 15 000 **g** for 3 to 5 min stratifies the cytoplasm and leaves the pronuclei clearly visible (Wall, Pursel, Hammer and Brinster, 1985). There is only one report of gene transfer in cattle (Biery *et al.*, 1988). Injected eggs were transferred to ewes as temporary recipients. Of the 829 injected eggs 175 had continued development to the morula stage (21%). These embryos were transferred to recipient heifers and 79 foetuses were recovered at autopsy at day 60 of pregnancy (10% of injected eggs). There were several different treatments, but overall, there were four transgenic foetuses from 829 injected eggs (0·48%). [In the most effective treatment there were two transgenic foetuses from 120 eggs injected (1·67%).] These observations are similar to those in sheep when 0·84% of injected eggs survived to become transgenic lambs (Simons *et al.*, 1988, and unpublished).

The mechanisms that govern incorporation of the DNA into a chromosome are not understood. In most cases, the injected DNA is linked into chains before incorporation. It has been suggested that the injection of liquid breaks the chromosomes and that the repair mechanisms inadvertently include the injected DNA as they repair the chromosome (Palmiter and Brinster, 1986). The site of incorporation is apparently random and evidence in mice suggests that in approximately 5 to 10% of cases the site is within an endogenous gene (reviewed by Palmiter and Brinster, 1986).

The present methods of gene transfer are extremely primitive. The frequency of gene transfer is low, in some cases essential endogenous genes are damaged and the control of gene expression is poor. In the future some of these limitations may be overcome by the development of methods for site-directed changes to genes, but, because these will have even lower success rates than current methods, their use will require the development of embryo stem cells.

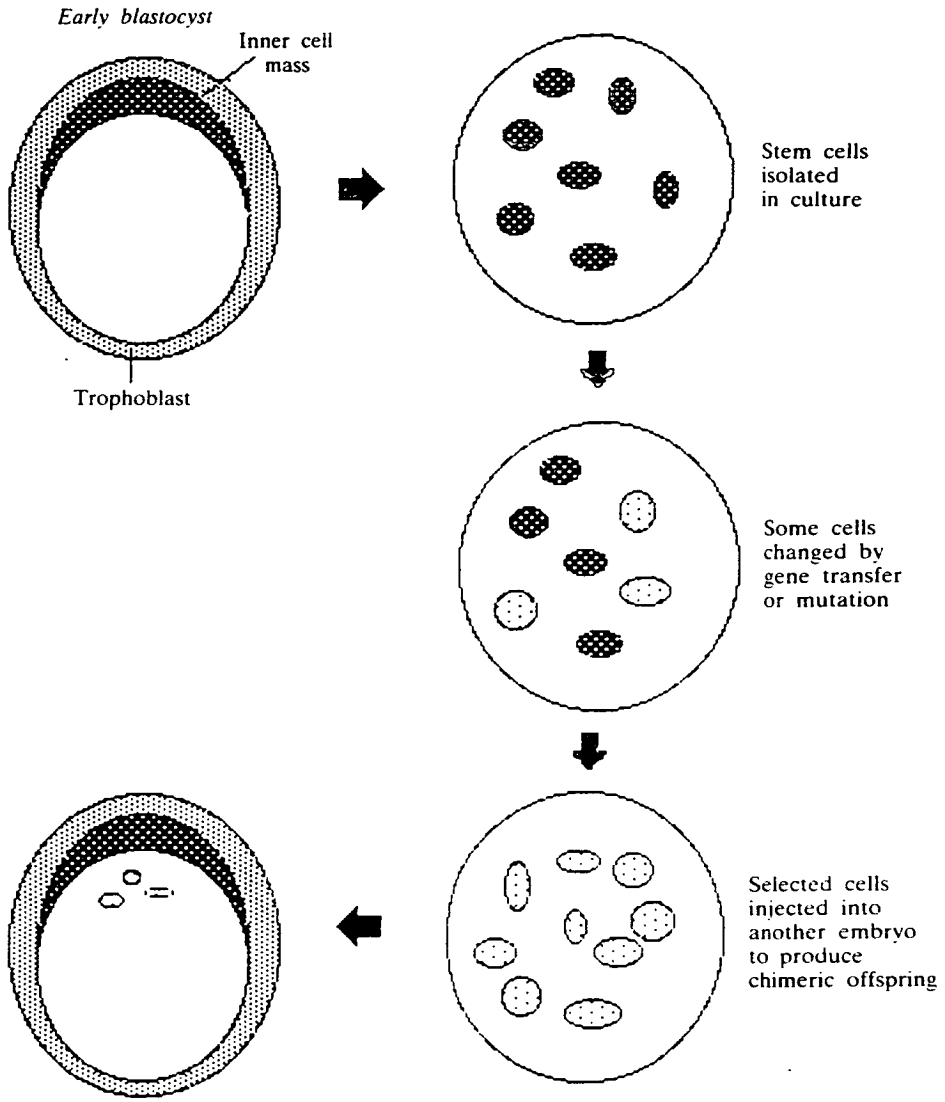


FIG. 6. Embryo stem cells. Embryo stem cells have been isolated from mouse embryos by culture in conditions that allow cell division, but inhibit differentiation. While in culture the cells may be modified by gene transfer or site directed mutation, before the transformed cells are selected. After transfer of selected cells into the cavity of another blastocyst some cells are able to colonize all of the tissues of the chimeric offspring, including the germline.

Embryo stem cells (Figure 6)

Embryo stem cells come from the inner cell mass of an embryo and retain the ability to divide in culture, but do not differentiate (Evans and Kaufman, 1981). In some cases, when injected into the blastocoel cavity of another embryo, they are capable of colonizing the foetus that develops from that embryo leading to the birth of a chimeric animal with tissues made up of cells from the recipient embryo and the stem cell (Bradley, Evans, Kaufman and Robertson, 1984). At present, stem cells have only been isolated from mouse embryos. Manipulation of murine embryo stem cells in culture before transfer, has led to the establishment of transgenic

lines (Lovell-Badge, 1987), and the introduction of site-directed mutations (Hooper, Hardy, Handyside, Hunter and Monk, 1987; Kuehn, Bradley, Robertson and Evans, 1987; Thomas and Capecchi, 1987).

Stem cell-like colonies have been isolated for short periods from sheep and pig embryos, but reliable methods for their long-term culture have not been described (Handyside, Hooper, Kaufman and Wilmut, 1987; Piedrahita, Anderson, Martin, BonDurant and Pashen, 1988). Several advantages would arise from the use of embryo stem cells when making genetic changes. Firstly, it would be possible to confirm that the change had been made before using the cell. Secondly, site-directed additions or mutations would be possible, which would ensure (i) that genes are inserted in a site where they will do no damage and where genes would be expressed to the full and (ii) make it possible to alter endogenous genes.

The recent transfer of nuclei from cells of the inner cell mass of cattle and sheep embryos (Smith and Wilmut, 1988; S. M. Willadsen, personal communication), suggests that, if embryo stem cells can be isolated in these species, it may be possible to use them as donors of nuclei for transfer to enucleated oocytes. In this case the offspring will not be chimeric as the embryo would develop from a single genotype. (By contrast, transfer of nuclei from cells of the inner cell mass or embryo stem cells is not possible in mice at present.)

As emphasized by Wilmut and Smith (1988), there would be many potential applications if it becomes possible to transfer nuclei from embryo stem cells in livestock. Embryo stem cells would provide a large number of nuclei for transfer making it possible to create very large clone families. In addition, transfer of a nucleus from a modified embryo stem cell would provide the most efficient possible means of making a change to the genomic DNA. This approach would have all of the advantages of embryo stem cell manipulation without the disadvantage of producing chimeric offspring.

The potential of nuclear transfer and embryo stem cells in making genetic changes and in creating clone families makes this a key area for further research.

PRACTICAL APPLICATION

Use of new procedures in MOET nucleus dairy herds

It is only since the work of Nicholas and Smith (1983) (revised by Woolliams and Smith, 1988) that it has been feasible to consider rates of progress within nucleus herds being competitive with those of progeny testing schemes. These papers put forward methods to use multiple ovulation, embryo recovery and transfer, in improvement schemes (termed MOET nucleus schemes).

MOET nucleus schemes obtain progress by increasing the reproductive rate of the female and placing greater emphasis on selection among female candidates than was previously possible. As a result each candidate for selection has an extended family that can include full-sibs and half-sibs, both maternal and paternal; and by combining the information on performance from all female relatives (including the candidate herself if a female with a lactation record) enough information on the candidate's breeding value can be obtained in a sufficiently short time that annual rates of genetic progress are comparable with those obtained from progeny testing. Nicholas and Smith (1983) put forward two schemes based on this principle; an adult scheme in which candidates, both male and female, are selected after the first lactation of the females has been completed, and a juvenile scheme which reduced the generation interval still further by selecting on pedigree alone. Although the adult scheme offers slower rates of genetic improvement, it also has lower inbreeding per annum (Woolliams, 1989b) and it is this scheme that has been undertaken commercially (McGuirk, 1989). For this reason in the following discussion implications of the new procedures will be given in the context of adult schemes although similar implications will also apply to juvenile schemes.

Inbreeding. One of the problems likely to be encountered in the operation of MOET schemes is the rate of inbreeding and the reduction in genetic variation that this brings.

With increased inbreeding, production would be expected to decline, reductions in reproductive performance could make the schemes difficult to operate and the prospects for future rates of progress would diminish with loss of genetic variation. High rates of inbreeding arise through the use of family information which, whilst increasing accuracy, also increases the frequency of co-selection of sibs. This problem is particularly acute for males in dairy MOET schemes since all full-sibs have identical information. Thus, ways of reducing co-selection are required. Nicholas and Smith (1983) suggested that only one male from each full-sibship should be selected. Such restrictions are helpful, but at a cost to the rate of progress expected, since the selection intensity is decreased. The rate of inbreeding that could be expected in schemes of commercial size (using the restriction given above) have been simulated and indicate a rate of 4·5 to 5% per generation, that is over 1% per annum in a closed herd (J. Ruane, personal communication). Unchecked high rates of inbreeding would require importation into the herd and, in doing so, reduce genetic merit.

Embryo sexing. The first procedure that would prove useful in MOET schemes is the introduction of embryo sexing. To obtain the number of full-sib females required (n_f) it would be expected that n_f males would be produced, but because of restrictions only one of these would be used for breeding. The other males would be surplus. If however embryos could be sexed prior to transfer no (in theory), or few (in practice), surplus males would be produced. Thus there is a potential saving of resources required for transfer by $(n_f - 1)/(2n_f)$. Whilst it could be argued that some of these resources would be required to ensure the production of at least one male, it is nevertheless useful to examine how resources freed might be re-used: more sires or more dams/sire could be used, or larger female full-sib families could be produced. The potential effect this has on progress and inbreeding is shown in Table 2. Producing larger female full-sib families (i.e. increasing n_f) could increase progress — but also inbreeding. If more dams were mated to each sire, then progress could be enhanced and inbreeding reduced. The use of more sires with the same number of dams per sire could significantly decrease inbreeding with 6% more progress than without sexing. Although the examples given in Table 2 involve slightly different resources (i.e. T, see METHODS FOR ASSESSING GENETIC PROGRESS AND INBREEDING), this variation does not affect the conclusions. In practice, combinations of these strategies could be adopted. Thus the use of embryo sexing could provide a flexible means of increasing the effective population size whilst maintaining or even increasing potential progress within a fixed resource.

Mating design and IVF. A further strategy

TABLE 2

The effect on genetic progress and inbreeding of redistributing resources saved by embryo sexing in a MOET nucleus scheme using a nested mating system

Scheme description					Resources required† (T, calves per generation)	Embryo sexing	ΔG (phenotypic s.d. per generation)	Measure of population size‡ (individuals)
n_s	n_d	n_d/n_s	n_f	n_m				
4	36	9	4	4	288	No	0·388	7·54
4	36	9	7	1	288	Yes	0·455	6·54
4	56	14	4	1	280	Yes	0·425	7·77
6	54	9	4	1	270	Yes	0·411	10·07

† The expected number of calves available for selection.

‡ See METHODS FOR ASSESSING GENETIC PROGRESS AND INBREEDING. As the measure of population size increases, inbreeding decreases.

has been proposed by Woolliams (1989b) in which changes to the mating system in MOET are made. Briefly, instead of mating a cow to only one bull, a different bull is used for each recovery of embryos. The number of offspring per bull and per cow are kept constant. In this way, accuracy of evaluation will be all but maintained whilst the production of smaller full-sib groupings will allow a greater selection intensity. As a result, as with embryo sexing, there will be an opportunity to increase effective population size or genetic progress or both. The extreme version of this mating system is a full factorial mating system in which each cow is ·ated to each bull.

A problem with such a mating system might be an increase in the generation interval arising from the need to carry out several collections of embryos. This would be the only reliable way of ensuring that each sire obtains progeny from each cow — mixed semen goes some way towards this but extreme results may occur (Beatty, Bennett, Hall, Hancock and Stewart, 1969; Beatty, Stewart, Spooner and Hancock, 1976). If however oocytes were obtained from follicles by laparoscopy and fertilized *in vitro* then the increase in generation interval might be avoided and the full benefits of factorial mating might be achieved. The possibility depends on the frequency and effectiveness of the recovery of oocytes by laparoscopy.

The benefits of IVF could add a further dimension to the design of breeding schemes. It may be possible to increase the selection ıtensity among cows (the most accurately evaluated sex in MOET) by decreasing the number bred from and by increasing the number of sires bred from to avoid increasing the full-sib family size. This would demand bigger maternal half-sib families and would require IVF to complete the design within a reasonable timescale. The potential benefits can be seen by comparing a $4m \times 36f$ design with a $12m \times 12f$ design, both requiring $T = 288$ calves. Table 3 shows there are substantial benefits both in progress and in inbreeding. However, to make the design feasible approximately 36 embryos would be needed and assuming a success rate of 0·7 embryos per oocyte (perhaps optimistically), this would require 50 oocytes to be obtained in as short a period as possible. The extent to which the procedures of oocyte recovery, IVF and embryo culture can improve to make this possible in a sufficiently short period is unclear. Nevertheless, the example shows the potential for improvement through this pathway.

Compatability of embryo sexing and factorial mating. It is notable that the two strategies of embryo sexing and factorial mating are not automatically compatible. Embryo sexing saves resources in the nested mating systems through a reduction in surplus full-sib males. In factorial mating systems whilst the ability to sex embryos helps in making full-sib families more nearly the desired sex ration, the design, if carried out perfectly, creates fewer surplus males. Thus, there are fewer resources freed for redistribution.

Cloning. Nicholas and Smith (1983) considered cloning/embryo splitting to be beneficial to rates of progress since clone families gave increased accuracy. However, Woolliams (1989a) pointed out that with fixed resources (defined by the number of transfers) a conflict arises between the size of the clone families and the selection pressure that can be applied. With nested mating designs the effect of cloning was then equivocal with gains in progress arising only when the resources required for cloning or splitting were made available at the expense of sire families, a strategy that would be likely to exacerbate the problems of inbreeding.

However, clone families not only have the same genotype with respect to performance but also with respect to breeding. Creating clone families of size two is therefore a way of doubling the reproductive rate of the animal, and so is a potential way of carrying through more effectively the advanced mating designs previously discussed in relation to IVF. Thus what was perceived by Woolliams (1989a) as a two-way conflict between selection pressure and accuracy through cloning becomes a three-way trade-off between these two and mating design.

Table 3 gives an indication of how this trade-off might be resolved. With $T = 288$

TABLE 3

The effect on genetic progress and inbreeding of increasing the maternal half-sib family size through changes in the reproductive capability of the cow, including cloning (n_c = 2) in MOET nucleus herds using factorial mating systems

Scheme description					Resources required† (*T*, calves per generation)	ΔG (phenotypic s.d. per generation)	Measure of population size‡ (individuals)
n_s	n_d	n_f	n_m	n_c			
4	36	1	1	1	288	0·444	7·41
12	12	1	1	1	288	0·493	8·55
8	9	1	1	2	288	0·485	7·65

† The expected number of calves available for selection.
‡ See METHODS FOR ASSESSING GENETIC PROGRESS AND INBREEDING. As the measure of population size increases, inbreeding decreases.

calves and clone families of size two, 144 distinct genotypes can be produced. With factorial mating this could be produced by a design of $8m \times 9f$. This design could give rates of progress only slightly less than the $12m \times 12f$, with inbreeding similar to that for nested mating systems with no embryo sexing (Table 2). It involves mating each individual cow to only four sires and so considerably reduces the possible penalties in generation interval.

Compatability of embryo sexing and cloning. It was shown earlier that with the full factorial design there is less scope for using resources more effectively by embryo sexing. When cloning or splitting is introduced this does not remain the case. There is no advantage to cloning males since their clones do not add information and the additional reroductive rate is not required. Thus by sexing embryos and redistributing the resources, further developments are possible. A comparison between a factorial design using both cloning and sexing and a nested design with sexing is shown in Figure 7. The graph shows that for the schemes considered, when compared at the same rate of inbreeding as the nested design, the factorial design has the potential for increasing the rate of progress by 8%. The value of mating design in combination with embryo sexing and cloning merits further study.

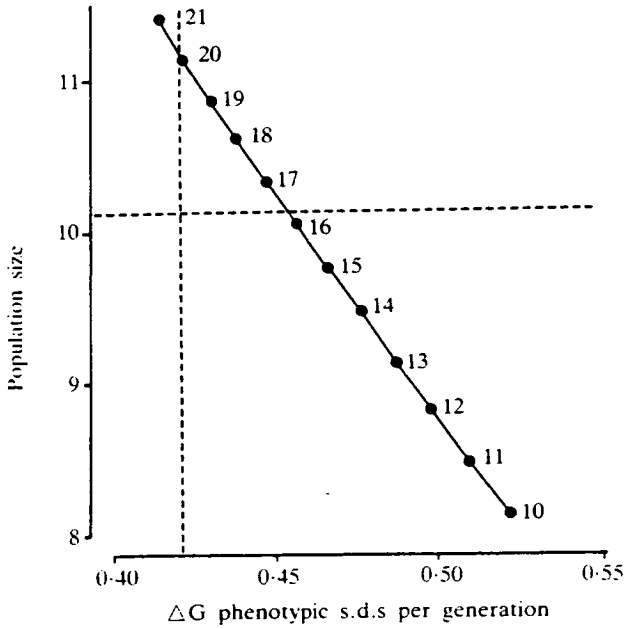


FIG. 7. Graph of changes in genetic progress and population size (hence inbreeding) in a factorial mating system (n_s = n_d = 10, n_f = n_m = 1, n_c = 2 for females, 1 for males; T = 300) by increasing the number of males and females selected on merit (from 10 to 21) from among which breeding stock are randomly selected. This is compared with a nested mating system using embryo sexing (n_s = 6, n_d = 10, n_f = 4, n_m = 1, T = 300) for which the maximal genetic progress and associated inbreeding are denoted by - - -.

Cytoplasmic inheritance. A further potential problem in the evaluation of males in MOET nucleus schemes is the reliance on maternal sibs (both full- and half-) which can introduce additional prediction errors. The size of the errors depend upon the degree to which cytoplasmic inheritance affects performance. This mode of inheritance is transmitted through the dam alone since the cytoplasm of an individual derives entirely from the oocyte from which it develops with no contribution from the sperm. Errors in evaluation arise because maternal sisters of a bull may benefit from this component and so influence selection, and although the bull has the identical component, he is unable to transmit it to his daughters. The extent of cytoplasmic variation as it affects performance is not known, but has been estimated to be from 0% (Kennedy, 1986) to 10% (Huizinga, Korver, McDaniel and Politiek, 1986) of total variation (cf. 25% for additive genetic variation).

This problem might be overcome by the procedure of nuclear transfer, if it could be refined to eliminate donor mitochondria. By transferring the nuclei of embryos in the breeding herd to enucleated oocytes from a single donor (or donors, if they originate from the same maternal line) evaluation would be more accurate through the removal of cytoplasmic variation. The transfer of nuclei has only to be carried out for the first generation of progeny born to each cow entering the herd. Although, within the herd, genetic progress (via nuclei) can be accurately measured, problems could arise in comparing the additive genetic merit of the MOET nucleus herd with others. This solution assumes that there are no interactions between inheritance via the cytoplasm and via the nucleus.

Use of new procedures in herds involved with progeny testing

The value of multiple ovulation, embryo recovery and embryo transfer (collectively abbreviated to ET, and involving more than the transfer of embryos) has been extensively discussed in the literature and has recently been reviewed by Ruane (1988).

The primary use of ET was identified by Hill and Land (1976) and Cunningham (1976) in increasing the selection intensity of cows used to breed bulls for progeny testing (the cb pathway). This is achieved by using multiple ovulation to increase the likelihood of obtaining a bull for testing from a particular cow and thereby reducing the number of cows required for producing bulls. The estimated increase in progress achievable depended on the accuracy with which superior cows were identified but varied between 2% and 10%. Although the production of bulls is the prime objective in the matings for the cb pathway, embryo sexing is of relatively minor value. Sexing would allow verification that male embryos had or had not been obtained at any recovery and thus whether or not further embryo recoveries were required. It would also enable only male embryos to be transferred (in sufficient numbers to be confident of obtaining at least one bull) and so unnecessary transfer costs could be avoided. However, any female embryos produced would be expected to be of high genetic merit and so would still be transferred if the costs of transfer were low enough. Thus, it is unlikely that embryo sexing would be in heavy demand for this purpose.

There are opportunities to use embryo sexing in the progeny test itself, although there are serious drawbacks. The objective in the progeny testing of bulls is to produce daughters. Thus the ability to sex embryos would provide ways of utilizing the limited testing resources more effectively, by avoiding the production of unwanted males. Fewer cows need be used in the herds involved in testing bulls, allowing more cows to be mated for other purposes. This flexibility could be used in a variety of ways: (i) reducing the costs of the scheme; (ii) ensuring more effective comparisons between bulls both within and across participating herds; and (iii) testing more bulls, thereby increasing the selection intensity for bulls to breed the next generation and increasing genetic progress as a consequence. Of the latter two possibilities the last has more potential for improving progress since bulls are already evaluated to a high accuracy. Nevertheless, the ability to

control more closely the distribution of daughters between herds would make the system more robust.

However, embryo sexing requires a source of embryos. Embryos could not be produced through the *in vitro* culture of oocytes that were obtained from slaughtered cows, since information on the dams performance is important for evaluation. The organization of embryo recoveries would therefore be a major undertaking for the testing agency. Furthermore, the use of multiple ovulation would introduce full-sibs into the testing procedure and prediction errors would increase through the common maternal inheritance of the full sisters (including cytoplasmic inheritance). The production of maternal half-sibs would reduce this problem but such a solution would add significantly to the organization required. Finally, multiple ovulation would also lead to a reduction in the number of cows producing daughters for testing. This creates a risk that the sampling base might become unrepresentative of the population as a whole and uncontrolled errors could occur through non-additive gene action. Thus, the apparent advantages of embryo sexing in progeny testing mask a serious problem in defining an adequate donor population that would not threaten its inherent robustness.

Use of new procedures in commercial dairy herds

The use of ET has been suggested for commercial herds. Assuming that 60% of the herd is used to breed replacements each year and that a single recovery of embryos could produce four calves, the use of ET would enable replacements to be bred from the top 15% of cows (as donors) and transferred to other herd members. Depending, as before, on the accuracy with which cows of high genetic merit are identified, the merit of the herd would be increased through the additional selection on the cows. With embryo sexing, only the female embryos need be transferred thus releasing an additional 30% of the herd for more economic beef production. However, McDaniel and Cassell (1981) and Van Vleck (1981) concluded that the gains from using ET alone in these circumstances would far from offset the costs of ET. It seems unlikely that embryo sexing, which also has a financial cost of its own, would alter this conclusion. Reductions in the cost of ET may eventually make this economically viable, but if so the impact would be widespread and require operation of ET on an immense scale (Hill and Land, 1976).

Benefits to commercial herds could arise from genetic lift using a procedure that could produce large clone families. The difference in the genetic merit of commercial herds and leading herds is equivalent to many years of genetic progress. The most profitable genotypes from the leading herds could be taken directly by cloning and transfer to the commercial herds *in toto* instead of the gradual process of upgrading by conventional (AI) breeding. For each herd, this would be an important opportunity and the dissemination of genetic merit throughout a population is an important application for the cloning. On a national scale, large numbers of clones would be produced cheaply from a few females of the very best breeding herds. As a result, the average genetic merit of the commercial herd could become greater than that of the breeding herds. A testing programme for achieving this and the potential benefits are considered by Nicholas and Smith (1983).

There are dangers that could arise from the resulting reduction in genetic variation. In practice, for a single herd, this could lead to losses in production through increased penetration of pathogenic and metabolic diseases. Moreover, progress beyond this level depends on the rate of improvement in the remaining breeding herds. The long-term prospect for improvement will depend on the effective population size of these herds. Since (i) the ability to define true effective sizes in populations under selection is still a matter for extensive theoretical research and (ii) the rate of mutation, which enables genetic variation to be maintained under selection, is also unknown, the minimum effective population size required cannot be determined. Embryo banks have been proposed to safeguard against extreme misfortunes. These would rely on freezing,

long-term storage and thawing of embryos as considered previously. However, the maintenance of this safeguard is also not without risk. It is still a matter for research to produce convincing strategies for the dissemination of genetic merit through cloning.

Implications of new procedures for beef breeding and production

The potential improvement in genetic progress for growth rate within MOET nucleus herds was first shown by Land and Hill (1975). More recently, improvement schemes for beef cattle incorporating this principle have been started in the United Kingdom. Although the principles behind the increase in progress are the same as for dairy schemes, differences do exist. Foremost amongst these is that males express the trait under selection, and so full-sib males have distinct indices of merit. The problems of inbreeding encountered in beef MOET schemes are thus less severe than in dairy schemes. Nevertheless, much of the discussion on the impact of new procedures on dairy MOET nucleus schemes would be of relevance to beef schemes. However detailed evaluations have yet to be made.

Over 80% of the beef in the European Economic Community is derived from the dairy herd and changes in the structure and management of the dairy herd have a major impact on beef production. IVM, IVF and embryo culture have already been incorporated into commercial schemes to improve beef production from dairy herds. Oocytes are recovered from ovaries obtained from slaughterhouses, matured *in vitro* and fertilized with semen from a beef bull *in vitro*. The embryos are then cultured and transferred to commercial dairy cows. By taking oocytes from crossbred heifers, a three-quarter-bred beef calf is produced to compare with the half-bred possible through conventional means. The use of embryo transfer also offers the opportunity for twinning. The embryos for transfer can be obtained cheaply but there is only minimal control of the genetic merit of the oocyte. The gain in numbers and quality will need to more than offset the cost of the embryos and their transfer compared with AI. If this system were to become widely used there could be a gradual grading up of the oocyte donors. Embryo sexing would be a beneficial addition since the male beef calf would prove more profitable. However, beef-type females will be required to provide oocytes.

If the three-quarter-bred calf is commercially viable then there would certainly be scope for the production of full beef calves. These could be provided in sufficient quantity if embryo stem cells and nuclear transfer could be combined to produce clones. When carried out on a large scale, the production of the embryos would be inexpensive. Unlike the previous example, full control over the genetic merit of the clone would be possible. Thus, crossbred calves could be produced combining desirable aspects of different beef breeds and utilizing heterosis in growth and health (Cundiff, 1970; Gregory, Cundiff, Koch, Laster and Smith, 1978; Gregory, Koch, Laster, Cundiff and Smith, 1978). The crossing of breeds with and without the double-muscling gene would produce heterozygotes which would combine the advantages of the gene in muscle growth with few of the problems encountered with the homozygote (Ménissier, 1982; Theissen and Rollins, 1982). The improvement in reliability and consistency in yield and eating qualities likely to result from clones would be of benefit to the industry (Meat and Livestock Commission, 1988). The opportunity would also exist for the specific advantages of the different breeds to be fully utilized. If the procedure for cloning were to be developed it would appear to have much potential.

Other implications of cloning

Genotypic and additive genetic variance. The variance between clone families estimates the total genotypic variance. This is the sum of the variation from all genetic components and is consequently greater than the additive genetic component. The variance is of intrinsic interest both practically and scientifically in defining more clearly the relative contributions of genotype and environment. The separation of genotype and environment has important scientific applications (e.g. Field, Woolliams, Dingwall

and Munro, 1984). By comparing the proportion of variance that is additive genetic (h^2) and genotypic (H^2) the proportion that is non-additive can be estimated. This would be of use in genetic evaluation where information from relatives was used since better approximations could be made to the covariance among relatives (by assuming a particular non-additive component such as dominance). Evaluation would improve through more appropriate modelling, but the impact would depend on the amount of non-additive variation and on the frequency of occurrence of relationships such as full-sibs for which non-additive genetic variation is an important contributor to the covariance. The computation involved in the introduction of non-additive genetic variation to methods using best linear unbiassed prediction was considered by Kennedy and Schaeffer (1988).

Genotype by environment interaction. There is continuing debate over the magnitude and importance of genotype by environment interactions and this review does not intend to further it. Nevertheless, it is likely that their existence will continue to be tested. When considering the size of interaction between breeds or strains the ability to produce clone families large enough to test each family in each environment would improve the efficiency with which such interactions are detected. Within populations, the situation is slightly different, in part due to the definition of the interaction. One definition uses the correlations of the 'genetic values' measured in the various environments (a correlation less than 1 indicating an interaction). With clones the 'genetic values' could be either genotypic values or, more conventionally, additive genetic (breeding) values. Clones give the opportunity for estimating the genotypic correlation with much fewer resources than those required for an estimate of the additive genetic correlation. However, the estimation of the additive genetic correlation need not be more efficient when using clones. Nevertheless, it is likely that the availability of clones will make the estimation of the genotypic correlation a requisite first step in any investigation of genotype by environment interactions.

Selection opportunities. A further use of cloning would be to create new more appropriate selection criteria. It may arise that heritable information that is of economic value can only be obtained from mutually exclusive environments. Currently, this information cannot be obtained without sib or progeny testing. However, the use of clones in each environment could overcome this problem and information from the members of the clone family could be combined to give a breeding index for the genotype. An extreme example of this may be found in the pig industry where important objectives are efficiency and eating quality. Without cloning, information on the latter precludes subsequent breeding. Such an application would als have considerable scientific value.

Central comparisons of isolated genotypes. Some breeding circumstances require isolation of breeding stock in several sites to avoid or reduce risks from disease. Cloning of the disparate stocks and transfer to a central testing site would allow effective comparison of both genotypes and environments.

Implications of gene transfer

Exploitation of gene transfer depends upon an ability to design genes that are only expressed in specific tissues, at particular stages of development and at desired levels. Whilst some proteins have 'housekeeping' functions and are required in all tissues and at all times, others are only made in specific tissues and at certain periods of development. The different mechanisms controlling these genes are unclear and gene transfer itself provides a powerful tool for research in thi area.

Early studies with the rat gene for the digestive enzyme elastase provide an example of the complexity of the controlling sequences and the power of gene transfer to increase understanding. This gene codes for a protease that is synthesized predominantly in the acinar cells of the pancreas. Typically there are 10 000 molecules of mRNA per cell in rats. Following transfer of the protein coding sequences, together with 205 base pairs upstream from the gene to mice, many made similar numbers of mRNA molecules (Swift, Hammer, MacDonald and Brinster, 1984). By contrast, there was very little expression in

other tissues. The same upstream sequence was also capable of directing synthesis of human growth hormone to murine acinar cells (Ornitz, Palmiter, Hammer, Brinster, Swift and MacDonald, 1985). As a sequence consisting of only the first 72 base pairs upstream was not effective, it seems that an important element of control of tissue specific expression lies in the 133 base pairs omitted. The observation that regulatory sequences from one gene can be used to direct synthesis of a protein to an unusual tissue created new opportunities in animal breeding.

Some applications of gene transfer depend upon the identification of single genes that have major, desirable effects. Very few genes that have major effects upon aspects of animal production have been identified in the present populations. These include the Booroola, Icelandic and Cambridge genes affecting ovulation rate in sheep (Piper and Bindon, 1985; Jonmundsson and Adalsteinsson, 1985; Hanrahan and Owen, 1985) and the 'halothane' gene in pigs that increases yield of lean meat, but also increases deaths as a result of stress (Webb, Southwood and Simpson, 1987). These genes have such large effects that they have been difficult to exploit practically. New breeding strategies have been advanced for selection to reduce the deleterious effects of such genes and for the identification of genes with rather smaller effects (Roberts and Smith, 1982).

There are three aspects of cattle production that are the subject of transgenic research at present. These are modification of milk composition, and improvement in the efficiency of lean meat and milk production. They will be considered in turn but it should be remembered that the procedures are very new and it is probable that many potential applications have yet to be conceived.

There are several ways in which it may be useful to modify milk composition. Transgenic mice carrying the sheep beta-lactoglobulin gene produced very large quantities of authentic sheep protein in their milk (Simons, McClenaghan and Clark, 1987). This observation suggests that copies of genomic DNA coding for milk protein genes are very likely to be expressed at high levels in a tissue specific manner. In this case it may be possible for human proteins, such as lactoferrin, to be secreted in cow's milk, making the milk a better substitute for human milk when given to babies.

One application of changing composition is to direct secretion of proteins that are not normally present in milk to the mammary gland. Transgenic sheep have been created that secrete human clotting Factor IX in their milk by including the coding sequences for the human protein within the beta-lactoglobulin gene (Clark, Ali, Archibald, Bessos, Brown, Harris, McClenaghan, Prowse, Simons, Whitelaw and Wilmut, 1989). The control sequences have directed secretion of the human liver protein to the mammary gland. The concentration of Factor IX in the milk of these sheep is low, but new genes are being tested. This observation shows that human proteins that are needed for the treatment of disease can be produced in the milk of transgenic farm animals. Factor IX is one of the proteins normally present in blood, that are required for clotting of blood. Haemophiliacs suffer from hereditary deficiencies of these proteins and must receive regular injections of replacement. The protein is currently prepared from human blood, but this is expensive and in the past the product has contained viruses responsible for hepatitis and AIDS. There are other potential sources of these proteins, but these also have drawbacks: genetically modified bacteria are not capable of carrying out the post-translational modifications made to these proteins in the liver; and mammalian cells in culture, whilst able to synthesize authentic proteins, only do so at low levels (Anson, Austin and Brownlee, 1985; Busby, Kumar, Joseph, Halfpap, Insley, Berkner, Kurachi and Woodbury, 1985). There may be proteins with industrial application that could be made by transgenic farm animals. It remains to be determined which proteins the mammary gland is capable of synthesizing, at what concentration and whether the cost will be lower than that for other systems of production.

Growth and lactation are both increased by administration of growth hormone by frequent injection and it may be useful to transfer genes containing growth hormone sequences

(reviewed by Wagner, 1985). Mice (Palmiter, Brinster, Hammer, Trumbauer, Rosenfeld, Birnberg and Evans, 1982), sheep and pigs (reviewed by Rexroad and Pursel, 1988) have been produced that synthesize growth hormone in the liver. Control sequences of the gene metallothionein were linked to the structural sequences for growth hormone. As this gene is not responsive to the feed-back mechanisms that normally govern growth hormone release, the concentrations of the hormone in the circulation is often higher than normal. In some instances this causes changes in metabolism heading to more rapid growth and reduction in fat content of the carcass (reviewed by Pursel, Miller, Bolt, Hammer, Palmiter and Brinster, 1988). However, in some pigs there were very serious side-effects, including reduction in libido in boars, anoestrus in females, arthritis and susceptibility to stress.

These observations confirm that it is critical to the exploitation of gene transfer that gene expression can be limited to specific stages of development. In future this level of control may be achieved by improvements in regulatory sequences including those that could be switched on by administered substances.

There is a great opportunity to use several procedures of embryo manipulation to reduce the considerable resources required for gene transfer. The process involves two stages: (i) obtaining intact incorporation and expression of the gene; and (ii) testing whether or not the transgenic animal is healthy and economically productive.

With current techniques this process requires two generations. In the first generation incorporation can be verified but only after the birth of the animal and, as mentioned previously, the proportion that are transgenic is small. Expression can also be verified in this generation. However, since each transgenic animal produced is a distinct line (differing from others by the number of copies incorporated, the sites of incorporation and levels of expression), a true evaluation of the performance cannot be made. Thus, a further generation is required in which offspring carrying the transgene of a particular transgenic line are compared with controls.

Using embryo stem cells and nuclear transfer a whole generation may be omitted. Verification of incorporation can be made before the transfer of the embryo, thus every animal born would be of value. Furthermore, in the first generation, two clone families could be produced that differ only in the transgene. This would enable initial testing to be carried out in the same generation.

A second benefit of using clone families from stem cells and nuclear transfer would be a reduction in the numbers needed to carry out the performance testing to a given accuracy. This arises because the underlying genotypic variation would be removed. Smith, Meuwissen and Gibson (1987) examined th number of progeny required to detect effects of various magnitudes without using clones. Using clone families the scale of the testing would be reduced by a factor $(1 - H^2)$ where H^2 is the broad heritability of the primary trait of interest. An implicit assumption is that there is no interaction between the transgene and its genetic background. Therefore, it is likely that further testing will still be required in a wider population. Nevertheless, the initial screening will provide a valuable decision tool for the expensive task of incorporating the new genetic variation into the livestock population. Smith *et al.* (1987) provides a very full account of the procedures and strategies for achieving this.

Finally, an alternative form of genetic manipulation is the induction of site-directed mutation. In this method DNA sequences are designed to replace or modify existing genes. This would broaden the opportunities to include the removal or modification of an existing gene product. Present procedures simply add an additional gene which functions in the presence of the animals complement of genes. These genes are very likely to influence the effect of the transgene, perhaps adversely. For example, it may be useful to stop production of some milk proteins when using transgenic aminals to make protein in their milk. Transfer of recessive genes to an animal with a dominant allelle would be without effect, unless the dominant allele can be switched off. Practical applications for techniques of site-directed mutation have not been

considered in detail because they are not yet available.

CONCLUSIONS

Multiple ovulation, embryo recovery and transfer will continue to play an important rôle in (i) the introduction of new breeds or lines into established populations and (ii) the breeding of bulls for progeny testing. Its routine introduction into nucleus breeding schemes is currently underway in several countries and is an important extension of its use.

Embryo sexing, whilst valuable in MOET nucleus schemes in enabling resources to be used more effectively, does not have the flexibility offered by appropriate mixtures of sexed semen. Thus, in progeny testing schemes the benefits of sex selection (Van Vleck, 1981) cannot be obtained by embryo sexing without overcoming major organizational problems.

Cloning has many applications although some depend heavily on the size of the clone family that can be produced. Possible applications include: (i) comparing environments and genotypes across environments; (ii) the testing of transgenics; (iii) the production of substantial improvements in the genetic progress of MOET nucleus schemes through greater and more flexible reproductive performance per genotype; and (iv) more efficient beef production in dairy herds.

Oocyte recovery and IVF has already found commercial application in beef production from dairy herds. It has further potential applications in MOET nucleus schemes, but its usefulness cannot be determined without further research.

Finally, three new techniques have the potential for making a major impact over the next decade and will provide the focus for much research. These are nuclear transfer, embryo stem cells and gene transfer. The importance of nuclear transfer depends on (i) the degree to which cytoplasmic inheritance is important in production, (ii) whether or not it can be combined with embryonic stem cells for the routine production of large clone families, and (iii) the impact of gene transfer. Embryo stem cells also have an important rôle in the conduct of gene transfer. The usefulness of gene transfer itself will depend heavily on the ability to control the expression of genes of major effect.

ACKNOWLEDGEMENTS

The authors are very grateful to Mr R. Thompson, Professor W. G. Hill, Dr B. McGuirk and Mr W. R. Christie for many helpful discussions and for their comments on the manuscript.

REFERENCES

ANSON, D. S., AUSTIN, D. E. G. and BROWNLEE, G. G. 1985. Expression of active human clotting factor IX from recombinant DNA clones in mammalian cells. *Nature, London* **315:** 683-685.

BEATTY, R. A., BENNETT, G. H., HALL, J. G., HANCOCK, J. L. and STEWART, D. L. 1969. An experiment with heterospermic insemination in cattle. *Journal of Reproduction and Fertility* **19:** 491-502.

BEATTY, R. A., STEWART, D. L., SPOONER, R. L. and HANCOCK, J. L. 1976. Evaluation by the heterospermic isemination technique of the differential effect of freezing at −196°C on fertility of individual bull semen. *Journal of Reproduction and Fertility* **47:** 377-379.

BIERY, K. A., BONDIOLI, K. R. and DEMAYO, F. J. 1988. Gene transfer by pronuclear injection in the bovine. *Theriogenology* **29:** 224 (Abstr.).

BRADLEY, A., EVANS, M., KAUFMAN, M. H. and ROBERTSON, E. J. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature, London* **309:** 255-256.

BRINSTER, R. L., CHEN, H. Y., TRUMBAUER, M. E., YAGLE, M. K. and PALMITER, R. D. 1985. Introducing genes into mice by direct injection. *Proceedings of the National Academy of Sciences* **82:** 4438-4442.

BUSBY, S., KUMAR, A., JOSEPH, M., HALFPAP, L., INSLEY, M., BERKNER, K., KURACHI, K. and WOODBURY, R. 1985. Expression of active human Factor IX in transected cells. *Nature, London* **316:** 271-273.

CHRISTIE, W. B. 1986. *Proceedings of the 5th Annual Conference of the American Society for Embryo Transfer*, pp. 33-44.

CLARK, A. J., ALI, S., ARCHIBALD, A. L., BESSOS, H., BROWN, P., HARRIS, S., MCCLENAGHAN, M., PROWSE, C., SIMONS, J. P., WHITELAW, C. B. A. and WILMUT, I. 1989. The molecular manipulation of milk composition. *Genome* In press.

CUNDIFF, L. V. 1970. Experimental results on crossbreeding cattle for beef production. *Journal of Animal Science* **30:** 694-705.

CUNNINGHAM, E. P. 1976. The use of egg transfer techniques in genetic improvement. In *Egg Transfer in Cattle* (ed. L. E. A. Rowson), pp. 345-354. Commission of the European Communities, Luxembourg.

ELLIS, S. B., BONDIOLI, K. R., WILLIAMS, M. E., PRYOR, J. H. and HARPOLD, M. M. 1988. Sex determination of bovine embryos using male-specific DNA probes. *Theriogenology* **29:** 242 (Abstr.).

EVANS, M. J. and KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature, London* **292:** 154-156.

EYESTONE, W. H., VIGNIERI, J. and FIRST, N. L. 1987. Co-culture of early bovine embryo with oviductal epithelium. *Theriogenology* **27:** 228 (Abstr.).

FIELD, A. C., WOOLLIAMS, J. A., DINGWALL, R. A. and MUNRO, C. S. 1984. Animal and dietary variation in the absorption and metabolism of phosphorus in sheep. *Journal of Agricultural Science, Cambridge* **103:** 283-291.

FREEMAN, A. E. 1988. Breeding programs in dairy cattle — current and future considerations. In *Advances in Animal Breeding* (ed. S. Korver, H. A. M. van der Steen, J. A. M. van Arendonk, H. Bakker, E. W. Brascamp and J. Dommerholt), pp. 73-89. Pudoc, Wageningen.

GORDON, I. 1982. Synchronisation of estrus and superovulation in cattle. In *Mammalian Egg Transfer* (ed. C. E. Adams), pp. 63-80. CRC Press, Boca Raton, Florida.

GOWE, R. S., ROBERTSON, A. and LATTER, B. D. H. 1959. Environment and poultry breeding problems. 5. The design of poultry control strains. *Poultry Science* **38:** 462-471.

GREGORY, K. E., CUNDIFF, L. V., KOCH, R. M., LASTER, D. B. and SMITH, G. M. 1978. Heterosis and breed maternal and transmitted effects in beef cattle. III. Growth traits in steers. *Journal of Animal Science* **47:** 1054-1062.

GREGORY, K. E., KOCH, R. M., LASTER, D. B., CUNDIFF, L. V. and SMITH, G. M. 1978. Heterosis and breed maternal and transmitted effects in beef cattle. I. Preweaning traits. *Journal of Animal Science* **47:** 1031-1041.

HAMMER, R. E., PURSEL, R. D., REXROAD, C. E., WALL, R. J., BOLT, R. J., EBERT, J., PALMITER, R. D. and BRINSTER, R. L. 1985. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature, London* **315:** 680-683.

HAMMOND, J. 1950. Problems concerning the transplantation of fertilised ova or "artificial pregnancy". *Anales de la Facultad de Medicina, Montevideo* **35:** 810-819.

HANDYSIDE, A., HOOPER, M. L., KAUFMAN, M. H. and WILMUT, I. 1987. Towards the isolation of embryonal stem cell lines from the sheep. *Roux's Archives of Developmental Biology* **196:** 185-190.

HANRAHAN, J. P. and OWEN, J. B. 1985. Variation and repeatability of ovulation rate in Cambridge ewes. *Animal Production* **40:** 529 (Abstr.).

HASLER, J. F., MCCAULEY, A. D., LATHROP, W. F. and FOOTE, R. H. 1987. Effect of donor-embryo-recipient interactions on pregnancy rate in a large-scale bovine embryo transfer program. *Theriogenology* **27:** 139-168.

HILL, W. G. and LAND, R. B. 1976. Superovulation and ovum transplantation in genetic improvement programmes. In *Egg Transfer in Cattle* (ed. L. E. A. Rowson), pp. 355-368. Commission of the European Communities, Luxembourg.

HOOPER, M., HARDY, K., HANDYSIDE, A., HUNTER, S. and MONK, M. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonisation by cultured cells. *Nature, London* **326:** 292-295.

HUIZINGA, H. A., KORVER, S., MCDANIEL, B. T. and POLITIEK, R. D. 1986. Maternal effects due to cytoplasmic inheritance in dairy cattle: influence on milk production and reproduction traits. *Livestock Production Science* **15:** 11-26.

HUNTER, R. H. F. 1984. Towards 100% fertilisation in inseminated cows, with particular reference to the site of sperm storage. *Animal Breeding Abstracts* **52:** 1-3.

JONES, K. W., SINGH, L. and EDWARDS, R. G. 1987. The use of probes for the Y chromosome in preimplantation embryo cells. *Human Reproduction* **2:** 439-445.

JONMUNDSSON, J. V. and ADALSTEINSSON, S. 1985. Single genes for fecundity in Icelandic sheep. In *Genetics of Sheep Reproduction* (ed. R. B. Land and D. W. Robinson), pp. 159-168. Butterworths, London.

KENNEDY, B. W. 1986. A further look at evidence for cytoplasmic inheritance of production traits in dairy cattle. *Journal of Dairy Science* **69:** 3100-3103.

KENNEDY, B. W. and SCHAEFFER, L. R. 1988. Reproductive technology and genetic evaluation. I. *Advances in Statistical Methods for Genet. Improvement of Livestock* (ed. D. Gionola and K. Hammond). Springer Verlag. In press.

KUEHN, M. R., BRADLEY, A., ROBERTSON, E. and EVANS, M. J. 1987. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations in mice. *Nature, London* **326:** 295-298.

LAMBERT, R. D., SIRARD, M. A., BERNARD, C., BELAND, R., RIOUX, J. E., LECLERC, P., MENARD, D. P. and BEDOYA, M. 1986. *In-vitro* fertilization of bovine oocytes matured *in vivo* and collected by laparoscopy. *Theriogenology* **25:** 117-133.

LAND, R. B. and HILL, W. G. 1975. The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Animal Production* **21:** 1-12.

LAWSON, R. A. S., ROWSON, L. E. A. and ADAMS, C. E. 1972. The development of cow eggs in the rabbit oviduct and their viability after re-transfer to heifers. *Journal of Reproduction and Fertility* **28:** 313-315.

LEIBO, S. P. 1988. Preservation of embryos. *Proceedings of 11th International Congress on Animal Reproduction and Artificial Insemination, Dublin*, Vol. 5, pp. 370-377.

LEIBO, S. P. and RALL, W. F. 1987. Increase in production of pregnancies by bisection of bovine embryos. *Theriogenology* **27:** 245 (Abstr.).

LOVELL-BADGE, R. H. 1987. Introduction of DNA into embryonic stem cells. In *Teratocarcinomas and Embryonic Stem Cells: a Practical Approach* (ed. E. J. Robertson), pp. 153-182. IRL Press, Oxford.

LU, K. H., GORDON, I., GALLAGHER, M. and MCGOVERN, H. 1987. Pregnancy established in cattle by transfer of embryos derived from *in vitro* fertilisation of oocytes matured *in vitro*. *Veterinary Record* **121:** 259-260.

LU, K. H., GORDON, I., MCGOVERN, H. and GALLAGHER, M. 1988. Production of cattle embryos by *in vitro* maturation and fertilisation of follicular oocytes and their subsequent culture *in vivo* in sheep. *Theriogenology* **29:** 272 (Abstr.).

MCDANIEL, B. T. and CASSELL, B. G. 1981. Effects of embryo transfer on genetic change in dairy cattle. *Journal of Dairy Science* **64:** 2484-2492.

McGuirk, B. 1989. The role of MOET schemes in developing countries. *Theriogenology*. In press.

Mauleon, P., Marina, J.-C., Benoit, M., Solari, A. and Chupin, D. 1970. Influence de differentes doses de PMSG et HCG, injectes en phase folliculaire du cycle oestrien sur le nombre et le rendement d'ovulations de vaches de race Francaise Frisonne Pie Noire. *Annales de Biologie Animale, Biochimie et Biophysique* **10:** *Suppl. 1*, 31-46.

Meat and Livestock Commission. 1988. *Corporate Plan 1988-1991*. p. 5. Meat and Livestock Commission, Bletchley.

Ménissier, F. 1982. General survey of the effect of double muscling on cattle performance. In *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production* (ed. J. W. B. King and F. Ménissier), pp. 23-53. Martinus Nijhoff, The Hague.

Newcomb, R. 1979. Surgical and non-surgical transfer of bovine embryos. *Veterinary Record* **105:** 432-434.

ewcomb, R. 1982. Egg recovery and transfer in cattle. In *Mammalian Egg Transfer* (ed. C. E. Adams), pp. 81-118. CRC Press, Boca Raton, Florida.

Newcomb, R., Christie, W. R. and Rowson, L. E. A. 1978. The non-surgical recovery and transfer of bovine embryos. In *Control of Reproduction in the Cow* (ed. J. M. Sreenan), pp. 292-304. Martinus Nijhoff, The Hague.

Nicholas, F. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo traansfer and splitting. *Animal Production* **36:** 341-353.

Ornitz, D. M., Palmiter, R. D., Hammer, R. E., Brinster, R. L., Swift, G. H. and MacDonald, R. J. 1985. Specific expression of an elastase-human growth hormone fusion gene in pancreatic acinar cells of transgenic mice. *Nature, London* **313:** 600-603.

Palmiter, R. D. and Brinster, R. L. 1986. Germline transformation of mice. *Annual Review of Genetics* **20:** 465-499.

Palmiter, R. D., Brinster, R. L., Hammer, R. E., Trumbauer, M. E., Rosenfeld, M. G., Birnberg, N. C. and Evans, R. M. 1982. Dramatic growth of mice that develop from eggs microinjected with matallothionein-growth hormone fusion genes. *Nature, London* **300:** 611-615.

arrish, J. J., Susko-Parrish, J. L., Leibfried-Rutledge, M. L., Critser, E. S., Eyestone, W. H. and First, N. L. 1986. Bovine *in vitro* fertilisation with frozen-thawed semen. *Theriogenology* **25:** 591-600.

Parrish, J. J., Susko-Parish, J. L., Winer, M. A. and First, N. L. 1988. Capacitation of bovine sperm by heparin. *Biology of Reproduction* **38:** 1171-1180.

Piedrahita, J. A., Anderson, G. B., Martin, G. R., BonDurant, R. H. and Pashen, R. L. 1988. Isolation of embryonic stem cell-like colonies from porcine embryos. *Theriogenology* **29:** 286 (Abstr.).

Piper, L. R. and Bindon, B. M. 1985. The single gene inheritance of the high litter size of the Booroola Merino. In *Genetics of Reproduction in Sheep* (ed. R. B. Land and D. W. Robinson), pp. 115-125. Butterworths, London.

Pursel, V. G., Miller, K. F., Bolt, D. J., Hammer, R. E., Palmiter, R. D. and Brinster, R. L. 1988. Insertion of growth hormone genes into pig embryos. *International Symposium on Biotechnology in Growth Regulation*. Butterworths, London. In Press.

Pursel, V. G., Miller, K. F., Pinhert, C. A., Palmiter, R. D. and Brinster, R. L. 1988. Effect of ovum cleavage stage at micromanipulation on embryonic survival and gene integration in pigs. *Proceedings of 11th International Congress on Animal Reproduction and Artificial Insemination, Dublin*. Vol. 4, p. 480.

Rajakoski, E. 1960. The ovarian follicular system in sexually mature heifers with special reference to seasonal, cyclical and left-right variations. *Acta Endocrinologica, Suppl. 52*.

Rall, W. F. and Fahy, G. M. 1985. Ice-free cryopreservation of mouse embryos at −19°C by vitrification. *Nature, London* **313:** 573-575.

Rall, W. F., Wood, M. J., Kirby, C. and Whittingham, D. G. 1987. Development of mouse embryos cryopreserved by vitrification. *Journal of Reproduction and Fertility* **80:** 499-504.

Renard, J. P., Heyman, Y. and Ozil, J. P. 1982. Congelation de l'embryon bovin. Une nouvelle method de decogelation pour le transfert cervical d'embryons conditionnes une seule fois en paillettes. *Annales de Medecine Veterinaire* **126:** 23-32.

Rexroad, C. E. and Pursel, V. G. 1988. Status of gene transfer in domestic animals. *Proceedings of 11th International Congress on Animal Reproduction and Artificial Insemination, Dublin*. Vol. 5, pp. 28-35.

Roberts, R. C. and Smith, C. 1982. Genes with large effects — theoretical aspects in livestock breeding. *Proceedings of 2nd World Congress on Genetics Applied to Livestock Production, Madrid*. Vol. 6, pp. 420-438.

Rossant, J. 1976. Post-implantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *Journal of Embryology and Experimental Morphology* **36:** 283-290.

Rowe, R. F., Del Campo, M. R., Critser, J. K. and Ginther, O. J. 1980. Embryo transfer in cattle: nonsurgical transfer. *American Journal of Veterinary Research* **41:** 1024-1028.

Rowson, L. E. A., Moor, R. M. and Lawson, R. A. S. 1969. Fertility following egg transfer in the cow: effect of method, medium and synchronisation of oestrus. *Journal of Reproduction and Fertility* **18:** 517-523.

Ruane, J. 1988. Review of the use of embryo transfer in the genetic improvement of dairy cattle. *Animal Breeding Abstracts* **56:** 437-446.

Saumande, J., Chupin, D., Mariana, J. C., Ortavant, R. and Mauleon, P. 1978. Factors affecting the variability of ovulation rates after PMSG stimulation. In *Control of Reproduction in the Cow* (ed. J. M. Sreenan), pp. 195-224. Martinus Nijhoff, The Hague.

Schams, D., Menzer, Ch., Schallenberger, E., Hoffmann, B., Hahn, J. and Hahn, R. 1978. Some studies on pregnant mare serum gonadotrophin (PMSG) and on endocrine responses after application for superovulation in cattle. In *Control of Reproduction in the Cow* (ed. J. M. Sreenan), pp. 122-143. Martinus Nijhoff, The Hague.

Simons, J. P., McClenaghan, M. and Clark, A. J.

1987. Alteration of the quality of milk by expression of sheep B-lactoglobulin in transgenic mice. *Nature, London* **328:** 530-532.

SIMONS, J. P., WILMUT, I., CLARK, A. J., ARCHIBALD, A. L., BISHOP, J. O. and LATHE, R. 1988. Gene transfer into sheep. *Biotechnology* **6:** 179-183.

SMITH, C., MEUWISSEN, T. H. E. and GIBSON, J. P. 1987. On the use of transgenes in livestock improvement. *Animal Breeding Abstracts* **55:** 1-10.

SMITH, L. C. and WILMUT, I. 1988. Factors influencing nuclear transplantation in sheep embryos. *Journal of Reproduction and Fertility Abstracts* **1:** 10 (Abstr.).

STAIGMILLER, R. B. and MOOR, R. M. 1984. Competence of ovine oocytes matured outside the follicle. *Gamete Research* **9:** 221-229.

SWIFT, G. H., HAMMER, R. E., MACDONALD, R. J. and BRINSTER, R. L. 1984. Tissue specific expression of the rat elastase 1 gene in transgenic mice. *Cell, USA* **38:** 639-646.

THEISSEN, R. B. and ROLLINS, W. C. 1982. A comparison of normal and heterozygous animals for double muscling in British breeds of cattle. In *Muscle Hypertrophy of Genetic Origin and Its Use to Improve Beef Production* (ed. J. W. B. King and F. Ménissier), pp. 54-69. Martinus Nijhoff, The Hague.

THOMAS, K. R. and CAPECCHI, M. R. 1987. Stem cells. *Cell, USA* **51:** 503-512.

TROUNSON, A. O., WILLADSEN, S. M., ROWSON, L. E. A. and NEWCOMB, R. 1976. The storage of cow eggs at room temperature and at low temperature. *Journal of Reproduction and Fertility* **46:** 173-178.

VAN VLECK, L. D. 1981. Potential genetic impact of artificial insemination, sex selection, embryo transfer, cloning and selfing in dairy cattle. In *New Technologies in Animal Breeding* (ed. B. G. Seidel and S. M. Seidel), pp. 221-242. Academic Press, London.

WAGNER, T. E. 1985. The role of gene transfer in animal agriculture and biotechnology. *Canadian Journal of Animal Science* **65:** 539-552.

WALL, R. J., PURSEL, V. G., HAMMER, R. E. and BRINSTER, R. L. 1985. Development of porcine ova that were centrifuged to permit visualization of pronuclei and nuclei. *Biology of Reproduction* **32:** 645-651.

WEBB, A. J., SOUTHWOOD, O. I. and SIMPSON, P. 1987. The halothane test in improving meat quality. In *Control of Meat Quality* (ed. P. V. Tarrent, G. Eikelenbloom and G. Monin), pp. 297-315. Martinus Nijhoff, The Hague.

WILKIE, T. M., BRINSTER, R. L. and PALMITER, R. D. 1986. Germline and somatic mosaicism in transgenic mice. *Developmental Biology* **118:** 9-18.

WILLADSEN, S. M. 1982. Micromanipulation of embryos of the large domestic species. In *Mammalian Egg Transfer* (ed. C. E. Adams), pp. 185-210. CRC Press, Boca Raton, Florida.

WILLADSEN, S. M. 1986. Nuclear transplantation in sheep embryos. *Nature, London* **320:** 63-65.

WILMUT, I. 1986. Crypreservation of mammalian eggs and embryos. In *Manipulation of Mammalian Development* (ed. R. B. L. Gwatkin), pp. 217-24 Plenum Press, New York.

WILMUT, I. and ROWSON, L. E. A. 1973. Experiments on the low temperature preservation of cow embryos. *Veterinary Record* **92:** 686-690.

WILMUT, I. and SMITH, L. C. 1988. Biotechnology and the bovine embryo: at present and in the future. In *Proceedings of the 4th Scientific Meeting of European Embryo Transfer Association* (ed. M. Thibier), pp. 19-31. Fondation Marcel Merieux, Lyons.

WOOLLIAMS, J. A. 1989a. The value of cloning in MOET nucleus breeding schemes for dairy cattle. *Animal Production* **48:** 31-35.

WOOLLIAMS, J. A. 1989b. Modifications to MOET nucleus schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production*. In press.

WOOLLIAMS, J. A. and SMITH, C. 1988. The value of indicator traits in the genetic improvement of dairy cattle. *Animal Production* **46:** 333-345.

WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics, USA* **16:** 97-159.

Paper 17

Anim. Prod. 1993, 56: 179-186

0003-3561/93/20770179$02.00

Decision rules and variance of response in breeding schemes

J. A. Woolliams[1] and T. H. E. Meuwissen[2]

[1]*Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH26 9PS*
[2]*Research Institute for Animal Production 'Schoonoord', PO Box 501, 3700 AM Zeist, The Netherlands*

Abstract

Selection decisions in breeding schemes can involve choices between candidates evaluated to different accuracies. A Bayesian framework is put forward for choosing among candidates, and it is shown that attaching loss functions for estimation errors makes this process different from selecting upon best linear unbiased predictions alone. Examples are given using both linear and quadratic loss to show that when estimation errors are penalized, the selection process tends to select more unrelated and more accurately evaluated individuals. In a dairy cattle breeding scheme response was only slightly lower than that from selection on expected breeding values but with a much reduced coefficient of variation. However, if prediction errors are preferred, with the hope of selecting individuals whose breeding value are higher than expected, extra genetic progress could be obtained by favouring the selection of individuals with low accuracy. This requires consideration of more than a single generation.

With discrete generations and equal measurements on candidates the decision framework was shown to be equivalent to a single quadratic restriction on the selection scores of parents in the previous generation.

A framework based on Bayes decision theory could be simply applied to produce a flexible means for producers to select according to their individual risk preferences.

Keywords: *Bayesian theory, breeding programmes, dairy cattle, selection.*

Introduction

When confronted by selection decisions it is usual to select on an unbiased prediction of breeding value such as that provided by best linear unbiased prediction (BLUP). In some situations candidates for selection are of a single age group with the same amount of information on each. However, with BLUP, selection may be made across different age groups and environments, and this is seen as a major advantage of the method. The older age groups often accumulate extra information through extra records on their own performance and records from relatives. Although in general, because of selection progress, the younger group contain animals genetically superior to the older group, this is complicated by the extra information for older animals since this allows their merit to be more precisely determined.

How then does a breeder arrive at a decision between candidates? For the breeder what is important is the true breeding value — irrespective of age. Unfortunately all he has at his disposal is an unbiased prediction of the breeding value together with an estimate of its error variance which may vary widely between candidates. With normal distributions, the estimation using BLUP can be cast in Bayesian terms since the estimate and the variance of estimation errors equal the mean and variance of the posterior distribution but what is required to complete this approach is a Bayes decision process (Ferguson, 1967) for selection. This will require a profit function to describe how much is gained by making a decision. As an example, in estimating a breeding value the profit function depends on the estimation error (the difference between the value chosen as the estimate and the true value), and the Bayes estimate is the value that maximizes the expected profit. When the infinitesimal model is assumed along with a symmetric profit function, known genetic parameters and a multivariate normal distribution of breeding values in the base generation, BLUP estimates are the Bayesian estimates of breeding values.

This paper takes a Bayesian approach to the selection process rather than taking a Bayesian approach to prediction alone, and shows that these coincide only in special cases. Further it then considers the form of Bayes selection decisions for both symmetric and asymmetric loss functions of prediction error in the context of minimizing risk in dairy breeding schemes, and finally considers decision rules for achieving pre-determined targets on genetic merit of offspring. Schneeberger, Freeman and Boehlje (1982) considered a similar problem for one round of selection and used a different approach but their solution is equivalent to a special case considered here.

Material and methods

Profit functions and their expectation

An empirical profit function for selection may be derived by the following means. Suppose there is a set of N candidates with true breeding value μ_i and predicted breeding values $\hat{\mu}_i$ for $i = 1, \ldots N$. Define the profit for individual i as a function of μ_i, the true breeding value.

$$P(i) = \hat{\mu}_i + f'(|\mu_i - \hat{\mu}_i|) + g'(\mu_i - \hat{\mu}_i) + h'(\hat{\mu}_i - \mu_i)$$

where $f'(x) = g'(x) = h'(x) = 0$ for $x \leq 0$. An element of loss may arise from all errors in prediction since this might produce uncertainty in planning and investment, thus $f'(x) \leq 0$ for $x > 0$. Underestimation (i.e. $\hat{\mu}_i < \mu_i$) may yield subsequent benefits, thus $g'(x) \geq 0$ for $x > 0$. Overestimation (i.e. $\mu_i < \hat{\mu}_i$) may introduce a further cost for restoring and correcting gains below those expected, thus $h'(x) \leq 0$ for $x > 0$. This can be simplified by defining $f = f' + h'$ and $g = g' - h'$. Then, $P(i) = \hat{\mu}_i + f(|\mu_i - \hat{\mu}_i|) + g(\mu_i - \hat{\mu}_i)$ with $f(x) \leq 0$ and $g(x) \geq 0$ for $x > 0$.

The functional forms of f and g can be derived empirically, but in this paper two commonly used forms will be considered: a linear loss function, with $f(x) = ax$ and $g(x) = bx$; and a quadratic loss function, with $f(x) = ax^2$ and $g(x) = bx^2$. For the linear loss function, the expectation of the profit function for individual i is

$$E[P(i)] = \tfrac{1}{2}E[\hat{\mu}_i - a|x| + bx|x>0] + \tfrac{1}{2}E[\hat{\mu}_i - a|x| + |x \leq 0]$$
$$= \hat{\mu}_i - 0{\cdot}798\, a\sigma_i + 0{\cdot}399\, b\sigma_i = \hat{\mu}_i - 0{\cdot}798\,(a - \tfrac{1}{2}b)\sigma_i$$

where σ_i^2 is the prediction error variance for individual i and 0·798 is the expectation value of the absolute value of a single random sample from a N(0,1) distribution (equal to the intensity of selection for $p = \frac{1}{2}$). A similar derivation for quadratic loss gives:

$$E[P(i)] = \hat{\mu}_i - \sigma_i^2(a - \tfrac{1}{2}b).$$

Bayes decision rules

In animal breeding, the breeder is faced with the decision of choosing n from N candidates. There are $D = N!/((N-n)!n!)$ distinct selection rules $\{d_j;\ j = 1,\ldots D\}$ each defined by d_j: 'choose candidates $\{S_j\} = \{i_1, \ldots i_n\}$, reject remainder $\{R_j\}$'. A Bayes decision rule maximizes the expected profit under the posterior distribution of the breeding values which is multivariate normal with a vector of means $\hat{\mu}_i$ and a variance/covariance matrix **V** (**V** contains prediction error variances and covariances). Henceforward, the profit will be a function of a selection rule denoted by $P(d_j)$.

For a single sex, the expected profit from d_j could be described in two ways. Firstly, the profit could maximize the difference between those selected and those rejected. This is equivalent to maximizing the selection differential. Let $\Delta_w = \mathbf{c}_\mathbf{w}^T\boldsymbol{\mu}$ where the ith element of $\mathbf{c}_\mathbf{w}$ is n^{-1} if i is in $\{S_j\}$ and $-(N-n)^{-1}$ if i is in $\{R_j\}$, and $\boldsymbol{\mu}$ is the vector of breeding values. The predicted error variance of contrast $\mathbf{c}_\mathbf{w}$ is $\sigma^2_{\Delta w} = \mathbf{c}_\mathbf{w}^T\mathbf{V}\mathbf{c}_\mathbf{w}$, and is the variance of Δ_w conditional on the observed data. Hence, $E[P(d_j)]$ is $\hat{\Delta}_w - 0{\cdot}798\sigma_{\Delta w}(a - \frac{1}{2}b)$ for linear, or $\hat{\Delta}_w - \sigma^2_{\Delta w}(a - \frac{1}{2}b)$ for quadratic loss. Alternatively, the profit could be maximized by considering only those selected; results can be expressed in the same form as before by defining $\Delta = \mathbf{c}^T\boldsymbol{\mu}$ where element i of $\mathbf{c}$ is n^{-1} if i is in $\{S_j\}$ and 0 otherwise.

The latter method has the natural interpretation that $\Delta = \Sigma_i c_i\mu_i$ is the contribution to the genetic level of the next generation. This can be generalized to define a decision rule d_c: 'let individual i contribute c_i to the next generation' with $c_i \geq 0$ and $\Sigma c_i = 1$. Such a decision rule would be at least as good, over one generation, as $c_i = n^{-1}$ for some $\{S_j\}$. A further advantage of the method in involving the contribution vector, $\mathbf{c}$, is that it can also be generalized to two sexes. Suppose n_m males were to be chosen from N_m candidates and n_f females from N_f candidates, then if i is of sex x, $c_i = (2n_x)^{-1}$ and 0 otherwise.

It should be noted that in general $\mathbf{c}_\mathbf{w}$ and $\mathbf{c}$ give different results. The difference between $\mathbf{c}$ and $\mathbf{c}_\mathbf{w}$ can be examined by considering

$$\mathrm{Var}(\Sigma c_i\mu_i) = \mathrm{Var}(\Sigma_{I_S} c_i\mu_i) + \mathrm{Var}(\Sigma_{I_R} c_i\mu_i) + 2\,\mathrm{cov}(\Sigma_{I_S} c_i\mu_i, \Sigma_{I_R} c_i\mu_i),$$

where c_i is element i of either $\mathbf{c}$ or $\mathbf{c}_\mathbf{w}$. For $\mathbf{c}$ the last two terms are zero since $c_i = 0$ for i in $\{R\}$. But for $\mathbf{c}_\mathbf{w}$ the population has negative contributions where relatives are split between $\{R\}$ and $\{S\}$, hence these terms *add* to the expectation of the profit function. Thus $\mathbf{c}_\mathbf{w}$ favours more selection within families and might be expected to redistribute selection across families compared with $\mathbf{c}$.

A Bayes decision rule is equivalent to maximizing the predicted breeding values of selected individuals when $a = \frac{1}{2}b$, or $\mathbf{V} = \sigma^2\mathbf{I}$. Further if $f(x) = a$, $g(x) = b$ for $x > 0$ the profit function is indifferent to the size of errors. The expected profit is then $\hat{\Delta} - (a - \frac{1}{2}b)$, and the decision rule is again based solely on the predicted breeding values $\hat{\mu}_i$.

With f and g linear, the decision rule is a percentile of the posterior distribution of Δ, i.e. $\hat{\Delta} - 0{\cdot}798(a - \frac{1}{2}b)\sigma_\Delta$ is maximized. For example, if $0{\cdot}798(a - \frac{1}{2}b) = 1$ then the decision rule is $\hat{\Delta} - \sigma_\Delta$ and from referring to standard tables of the normal distribution this implies selection for the 16th percentile. If $a > \frac{1}{2}b$ the percentile is less than 50, but when $a < \frac{1}{2}b$ indicating that large benefits are obtained from positive errors the percentile chosen is greater than 50.

Examples

Example 1

Simulation was used to examine the difference between $\mathbf{c}$ and $\mathbf{c}_w$ when these decision rules were applied to a small, previously unselected population within a single generation. The population consisted of two half-sib families each of size 16, each split into two equal full-sib families with equal numbers of males and females. Selection indices for each individual were drawn from a normal distribution of mean 0 and s.d. 0·45, and with correlations of 0·45 and 1 between half and full-sibs, respectively. Vectors $\mathbf{c}$ and $\mathbf{c}_w$ were considered for the problem of selecting four males on their index while females remained unselected. Various values for $\alpha = a - \frac{1}{2}b$ were considered. The results for the sib selection of males is shown in Table 1. For low values of α, two randomly subsampled selected males were twice as likely to be unrelated if the decision rule $\mathbf{c}_w$ had been used rather than $\mathbf{c}$. Results were obtained from both linear and quadratic loss functions but since results were qualitatively similar only those for quadratic loss are presented in Table 1.

Table 1 *A comparison of decision rules on sib selection: $\mathbf{c}_w$ is a decision rule which promotes precision in estimating gains from choosing among candidates whereas* $\mathbf{c}$ *promotes precision in estimating the breeding value in the next generation*†

	α							
	0·5		1·0		2·0		4·0	
	$\mathbf{c}_w$	$\mathbf{c}$	$\mathbf{c}_w$	$\mathbf{c}$	$\mathbf{c}_w$	$\mathbf{c}$	$\mathbf{c}_w$	$\mathbf{c}$
P_{FS}	0·70	0·84	0·47	0·69	0·30	0·50	0·21	0·32
P_{HS}	0·09	0·05	0·14	0·09	0·16	0·15	0·18	0·16
P_U	0·20	0·11	0·39	0·22	0·54	0·35	0·61	0·52

† Results are from 400 simulated selection procedures (see text) and sib selection is measured by the probability that two randomly drawn selected males are full- or half-sibs or unrelated (P_{FS}, P_{HS} and P_U respectively). Standard errors from simulation are 0·05 $[\hat{P}(1-\hat{P})]^{\frac{1}{2}}$.

Example 2

The decision rule based on $\mathbf{c}$ using a quadratic loss function was applied to the first generation of selection of the adult MOET (multiple ovulation and embryo transfer) scheme (IU) described by Woolliams (1989). The scheme was comprised of four males mated hierarchically to 36 females. Each female produced four offspring of each sex. The trait considered was assumed to have a heritability of 0·25; correlations between indices and breeding values were 0·56 and 0·66 for males and females respectively. Correlations between indices of relatives are given by Woolliams (1989).

Selection was for both sexes. Whilst the solution of the selection problem could be obtained by quadratic programming, approximate solutions were obtained by the following iterative method. Let $\hat{\boldsymbol{\mu}}^{(0)}$ be the vector of posterior expectations of breeding values and $\mathbf{V}$ be the variance/covariance matrix for the prediction errors. Predicted breeding values of all individuals were adjusted for their prediction error i.e. $\hat{\mu}_j^{(1)} = \hat{\mu}_j^{(0)} - \alpha c_j \mathbf{V}_{jj}$ where $c_j = n_m^{-1}$ for males and n_f^{-1} for females. An initial individual i_1 was selected by choosing i_1 to maximize $c_{i1}\hat{\mu}_{i1}^{(1)}$, which is equivalent to choosing the highest corrected value. After selection, breeding values for the remaining individuals were adjusted for their covariance with i_1 i.e. $\hat{\mu}_j^{(2)} = \hat{\mu}_j^{(1)} - 2\alpha c_{i1}\mathbf{V}_{i1j}$ and i_2 was selected to maximize $c_{i2}\hat{\mu}_{i2}^{(2)}$. In this way the required number of males and females were selected in the order arising from application of the algorithm (i.e. it was not obligatory to select all males first followed by females or *vice versa*). At the completion of the selection

$$\sum_{j=1}^{n_m + n_f} c_{ij}\hat{\mu}_{ij}^{(j)}$$

is the value of the profit function for the selection of the n individuals. Results from the simulations are presented in Table 2.

As the penalty arising from prediction error variance increases, selection increasingly avoids full-sibs. Both gain and variance are reduced as a result, the variance more clearly so. The restricted selection scheme which avoids full-sibs males described by Nicholas and Smith (1983), denoted IR by Woolliams (1989) is included for comparison. More gain was obtained with a smaller prediction error using the decision rule. In the context of teams of bulls (Genus, 1991), as α increased so selection moved away from related to unrelated teams.

Table 2 *The relationship between sib selection, gain and prediction error variance of future gain in an adult MOET scheme (with four males, 36 females and eight offspring per female) when a cost (α) is placed on prediction error†*

Cost (α)	P_{FS}	P_{HS}	Gain (phenotypic s.d./generation)	Prediction error variance
0	1	0	0·452(0·0074)	0·0268(0·0016)
1	0·73	0·17	0·458(0·0071)	0·0247(0·0015)
2	0·53	0·22	0·437(0·0075)	0·0227(0·0016)
4	0·29	0·29	0·437(0·0076)	0·0198(0·0014)
8	0·10	0·40	0·402(0·0076)	0·0178(0·0009)
Scheme IR‡	0	0·56	0·391	0·0191

† Results are from 400 simulated selection procedures (see text) and sib selection is measured by the probability that two randomly drawn selected males are full- or half-sibs (P_{FS} and P_{HS} respectively). Standard errors from simulation are in parentheses or are given by $0{\cdot}05[\hat{P}(1-\hat{P})]^{\frac{1}{2}}$.
‡ Results from Woolliams (1989) in which selection is restricted by no full-sib males.

Example 3

The previous examples have examined the use of the Bayesian selection procedure in conditions in which each individual of a particular sex has equal prediction error variance. As suggested in the **Introduction,** Bayes decision rules could play an important rôle in selection decisions when accuracies of evaluation, and therefore prediction errors, vary. To examine the consequences of applying decision rules to a population including different accuracies, they were applied to the deterministic model of open nucleus schemes developed by Meuwissen (1991). The first example considered selection rules based on a linear loss function and prediction error covariances were neglected. Thus selection was on a percentile of the distribution of an individual's prediction (see **Material and methods**) considered independently from all other selection choices. The percentiles chosen were 5%, 20%, 50% and 80%. The last decision rule rewarded risk-taking by attaching a positive weight to errors i.e. rewards from underestimates outweigh losses from errors *per se*, or $b>2a$. The scheme modelled consisted of a nucleus with 16 sires mated hierarchically to 64 donors and four offspring of each sex per donor. Male offspring were progeny tested in the base. The results are shown in Table 3.

Table 3 *The expected annual genetic gain (ΔG) and its standard deviation ($\sigma_{\Delta G}$) from percentile selection in an open nucleus scheme with progeny testing*

Percentile	E(ΔG) (σg†/year)	$\sigma_{\Delta G}$ (σg)	CV	Generation interval for sires to breed sires (year)
80	0·2859	0·0671	0·23	2·0
50	0·2853	0·0518	0·18	2·5
20	0·2640	0·0235	0·09	5·0
5	0·2418	0·0161	0·07	6·2

† Standard deviation of breeding values at equilibrium.

As the percentile increased from 5 to 80% so did the expected genetic progress and its standard deviation. The increase in the latter was most dramatic. For the step from the 20%-ile to 50%-ile progress increased 1·08-fold but the CV increased two-fold. A change in structure of the population took place as the value of the percentile increased, with a decrease in the number of progeny tested sires used to breed sires and consequently a reduction in generation interval. An important point to note is that the expected progress increases above that from selecting on expected breeding value alone (i.e. 50%-ile) if the breeder takes risks by selecting on the 80%-ile i.e. by rewarding uncertainty in prediction.

Example 4

A final example was examined in which Bayes decision rules with quadratic loss were used for the selection of males in an open nucleus scheme. In this example relationships between males were accounted for using the iterative scheme outlined in example 2. The parameters of the open nucleus scheme are the same as those for example 3. Results are shown in Table 4. The general results are very similar to the results for example 3. Progress is reduced by penalizing prediction error but the precision of estimating response is greatly improved.

Relationships with other restricted selection techniques

The selection decision rule $\mathbf{c}_Q^*$, for quadratic loss, maximizes $\mathbf{c}^T\hat{\boldsymbol{\mu}}-\alpha\mathbf{c}^T\mathbf{V}\mathbf{c}$ for a given α. At the solution $\mathbf{c}_Q^*$ the decision rule has variance $\beta = \mathbf{c}^{T*}_Q\mathbf{V}\mathbf{c}_Q^*$ and consequently maximizes $\mathbf{c}^T\hat{\boldsymbol{\mu}}$ for all admissible $\mathbf{c}$ under the constraint $\mathbf{c}^T\mathbf{V}\mathbf{c} = \beta$. In this sense α acts as a Lagrangian multiplier to solve the constrained maximization. Conversely, from theory on constrained maximization, if an admissible solution exists for $\mathbf{c}^T\mathbf{V}\mathbf{c} = \beta$ for some β, then the solution of the

Table 4 *The expected annual genetic gain (ΔG) and its standard deviation ($\sigma_{\Delta G}$) from an open nucleus scheme using a quadratic loss function in its selection of males*

Cost (α)	E(ΔG)	$\sigma_{\Delta G}$	CV	Generation interval for sires to breed sires (year)
0	0·2885	0·0606	0·21	2·5
5	0·2772	0·0346	0·12	3·8
10	0·2625	0·0247	0·09	5·1

constrained maximization of $\mathbf{c}^T\hat{\boldsymbol{\mu}} - \lambda\mathbf{c}^T\mathbf{V}\mathbf{c}$ for some value of λ. Thus for each α there is a correspondence between the problems of maximizing $\mathbf{c}^T\hat{\boldsymbol{\mu}} - \alpha\mathbf{c}\mathbf{V}\mathbf{c}$ and $\mathbf{c}^T\hat{\boldsymbol{\mu}}$ with $\mathbf{c}^T\mathbf{V}\mathbf{c} = \beta$.

For a linear loss function $\mathbf{c}^T\hat{\boldsymbol{\mu}} - \alpha'\sqrt{(\mathbf{c}^T\mathbf{V}\mathbf{c})}$ is maximized or, equivalently, $\mathbf{c}^T\hat{\boldsymbol{\mu}}$ for $\sqrt{\mathbf{c}^T\mathbf{V}\mathbf{c}} = \beta'$ for some β'; since $\mathbf{c}^T\mathbf{V}\mathbf{c}$ is positive, this in turn is equivalent to maximization with the constraint $\mathbf{c}^T\mathbf{V}\mathbf{c} = (\beta')^2$ and, is therefore equivalent to the unconstrained maximization $\mathbf{c}^T\hat{\boldsymbol{\mu}} - \alpha''\mathbf{c}^T\mathbf{V}\mathbf{c}$ for some α''. Since the converse argument also holds, the set of optimum decision rules obtained by varying α' for linear loss are the identical set obtained by varying α'' for quadratic loss.

Consider a population maintained and selected in a single environment with discrete generations where all individuals of the same sex have the same records on performance (though possibly different between the sexes) and in which n_m and n_f individuals are selected each generation. Such populations might arise in MOET nucleus breeding schemes for example. Woolliams and Thompson in the appendix of Woolliams (1989) showed that for candidates for selection V can be decomposed into the form

$$\mathbf{V} = \mathbf{X}_{OO}^{-1} + \mathbf{X}_{OO}^{-1}(\mathbf{X}_{MO}^T\mathbf{V}_{MM}\mathbf{X}_{MO} + 2\mathbf{X}_{MO}^T\mathbf{V}_{MF}\mathbf{X}_{FO} + \mathbf{X}_{FO}^T\mathbf{V}_{FF}\mathbf{X}_{FO})\mathbf{X}_{OO}^{-1}.$$

In this decomposition $\mathbf{X}_{OO}$ is a diagonal matrix with rows and columns corresponding to candidates for selection, $\mathbf{X}_{MO}$ and $\mathbf{X}_{FO}$ are $-h^{-2}$ times the incidence matrices relating sires and dams to offspring, and $\mathbf{V}_{YZ}$ are variance/covariance matrices for parents of sexes y and z. The diagonal element corresponding to a candidate of sex y in $\mathbf{X}_{OO}$ is $a_y^{-1} = (2h^{-2}) + \delta_y$ where δ_y is the reciprocal of the environmental variance if sex x has records and 0 otherwise. Restricting $\mathbf{c}^T\mathbf{V}\mathbf{c} \leq \beta$ is equivalent to

$$\mathbf{c}^T\mathbf{X}_{OO}^{-1}\mathbf{c} + \mathbf{c}^T\mathbf{X}_{OO}^{-1}(\mathbf{X}_{MO}^T\mathbf{V}_{MM}\mathbf{X}_{MO} + 2\mathbf{X}_{MO}^T\mathbf{V}_{MF}\mathbf{X}_{FO} + \mathbf{X}_{FO}^T\mathbf{V}_{FF}\mathbf{X}_{FO})\mathbf{X}_{OO}^{-1}\mathbf{c} \leq \beta.$$

For the contribution vector $\mathbf{c}$, $\mathbf{c}^T\mathbf{X}_{OO}^{-1}\mathbf{c} = (a_m n_m^{-1} + a_f n_f^{-1})$ and $\mathbf{c}^T\mathbf{X}_{OO}^{-1}\mathbf{X}_{YO}^T = a_m n_m^{-1}\mathbf{S}_{(ym)} + a_f n_f^{-1}\mathbf{S}_{(yf)} = \mathbf{S}_y$ where $\mathbf{S}_{(yz)}$ is the vector of selection scores for the parents of sex y due to selection of their offspring of sex z. Thus $\mathbf{S}_y$ is a selection score weighted between the sexes, with weights in favour of the sex from which individuals make the largest contribution (i.e. small n_y), and the least individual information (i.e. small δ_y). Thus restricting $\mathbf{c}^T\mathbf{V}\mathbf{c} \leq \beta$ is selecting according to the rule

$$\mathbf{S}_m^T\mathbf{V}_{MM}\mathbf{S}_m + 2\mathbf{S}_m^T\mathbf{V}_{MF}\mathbf{S}_f + \mathbf{S}_f^T\mathbf{V}_{FF}\mathbf{S}_f \leq \beta'$$
$$\text{where } \beta' = \beta - (a_m n_m^{-1} + a_f n_f^{-1}).$$

The Bayes decision rules for linear and quadratic loss are in these conditions therefore equivalent to selection under the single restriction imposed by a quadratic form of the selection score vectors. This is distinct from the common method of restriction by imposing the set of linear constraints on the selection score for each individual i of sex m and f in the parental generation. The latter condition is imposed for sire selection by Nicholas and Smith (1983).

Profit functions based on thresholds

Up to this point the profit functions have provided a means of formalizing the trade-off between progress and predictability. Some circumstances (e.g. those faced by artificial insemination cooperatives and nucleus breeding herds) may be best modelled by considering success or failure relative to a threshold, T. In this case the profit function for selecting individual i is of the form

$$P(i) = \begin{cases} af(\mu_i - T) & \text{for } \mu_i \geq T,\ a > 0 \\ bf(\mu_i - T) & \text{for } \mu_i < T,\ b \geq 0 \end{cases}$$

For the selection of a group $P(i)$ would be replaced by $P(d_i)$ for the ith decision rule, μ_i by Δ_i.

When $f(x) = \text{sign}(x)$ for all x (i.e. $f(x) = +1$ or -1 depending on whether x is positive or negative), then for an individual i the expected profit is:

$E(P(i)) = aP(\mu_i \geq T) - bP(\mu_i < T) = a - (a + b)\Phi(\theta_i)$ where $\theta_i = (T - \hat{\mu}_i)\sigma_i^{-1}$ and $\Phi(\cdot)$ is the cumulative standard normal distribution function. When $b < a$, this reflects the situation where an individual (e.g. a bull) yields more profit if its true breeding value exceeds the threshold T. For a group of individuals $P(i)$, $\hat{\mu}_i$ and σ_i are replaced by $P(d_i)$, $\hat{\Delta}_i$ and $\sigma_{\Delta i}$ respectively. Selection on this profit function has the properties: (i) maximizes the probability of achieving the threshold and selection always depends on σ_i for all values of a and b; (ii) for two individuals (groups) with equal $\hat{\mu}(\hat{\Delta})$, then if $\hat{\mu} > T$ the individual (group) with smaller error variance is selected and if $\hat{\mu} < T$ the individual (group) with greater error variance is selected; and (iii) profit is greater than zero if $T - \sigma_i\Phi^{-1}(a(a + b)^{-1}) < \hat{\mu}$ where $\Phi^{-1}$is the inverse function of Φ.

If $f(x) = x$ for all x, then profit increases with increasing breeding value but at a different rate above and below the threshold. The expected profit of individual i is: $E(P(i)) = \sigma_i[(a - b)(\varphi(\theta_i) + \theta_i\Phi(\theta_i)) - a\theta_i]$, where φ is the probability density function of N(0,1). If $a = b$, selection is on $\hat{\mu}_i$, i.e. equivalent to BLUP; if two individuals i, j have ($\hat{\theta}_i = \hat{\theta}_j$, $\sigma_i < \sigma_j$), individual j with the greater error variance is selected if the expected profit is positive and *vice versa*. As a consequence there will be some decisions for which $f(x) = \text{sign}(x)$ favours smaller error

variances but for which $f(x) = x$ favours the reverse. With both profit functions if T is chosen to be in advance of the predicted breeding values of individuals available and $a > b$ then greater risk is favoured.

Discussion

Nicholas (1989) commented that unless animal breeders produced breeding schemes that could deliver genetic progress over a decade with a coefficient of variation less than 0·2 then both their future and that of their schemes would be limited. This illustrates the importance of both high genetic gain and low variance of response. Meuwissen (1991) showed that optimizing breeding schemes for dairy cattle led to the use of young untested bulls and whilst this produced faster genetic progress it was also accompanied by extra variance in response.

Application of theory

The method adopted in this paper has been to construct Bayes decision rules using profit functions which attach costs to errors in prediction. This method seems consistent with the increasing use of BLUP as a means of evaluation since BLUP with its inclusion of prior distributional knowledge of genetic effects can be viewed in a Bayesian context. When candidates for selection had differing amounts of information, selection on the estimated breeding value (EBV) alone was seen to be equivalent to using Bayes profit functions that attached importance only to the errors but not its magnitude, as might be expected from the unbiased nature of the EBV. When linear or quadratic profit functions were introduced then selection was based not on the BLUP estimate but on the estimate plus a percentile of the error distribution (for linear profit) or a multiple of the error variance (for quadratic loss). This demonstrates that the Bayesian approach to the selection process is distinct from the estimation process since with all symmetric loss functions the Bayesian estimate of breeding value *per se* will remain the EBV. Meuwissen (1990 and 1991) adopted a quadratic utility function to choose between alternative breeding schemes, and this has some similarity to the one adopted here, but whereas the utility function was applied post-analysis and post-hoc to an entire breeding scheme in which all participants shared the same utility values, in this approach decisions rules are used throughout and participants can act individually.

The results show that for large population sizes the greatest rate of progress may not result from the selection of those individuals with the highest expected breeding values but by selection that encourages risk through rewarding prediction error variance. Selection on EBV maximizes the expected breeding values of the next generation of offspring (Goffinet and Elsen, 1984), but when considered over more than one generation this depends critically on whether all candidates have the same amount of information at the time of selection (Woolliams, 1990). In circumstances where error variances differ it was shown that a small advantage was obtained from using high percentiles as the selection criterion, although the additional risk variance of response was substantial. The logic for this is that the merit of the next but one generation depends not on the average of the offspring produced in the next generation but on the merit of those that will be selected from it. Thus two candidates with equal predicted breeding values and different prediction error variances will have offspring with the same predicted breeding values, but the candidate with the higher prediction error variance has a higher probability of producing an extreme offspring, which could be selected in the next generation (Woolliams, 1990). Meuwissen (1991) demonstrated a further example of this phenomenon in a study which showed that in the absence of inbreeding, selection upon EBV in closed nucleus schemes had a faster rate of progress than open schemes: in the open schemes candidates are selected from outside the nucleus but only when sufficient information accumulates to overcome their poorer pedigree information and these individuals take the place of candidates from within the nucleus which are less accurately evaluated. However, in all these examples the gain in progress was small.

The control of variance has some effect on inbreeding but only to a degree determined by the magnitude of the prediction error variance. The examples showed that penalizing error variance led to a reduction in the co-selection of full- and half-sibs, and that in discrete generations it was equivalent to a quadratic restriction on the selection scores of parents. The similar appearance can be misleading. If the error variance of parents is zero, the breeding scheme becomes completely predictable (Mendelian sampling apart) and selection would be upon EBV; and depending on the form of selection procedure, increased co-selection of relatives could still occur and contribute to increased rates of inbreeding above those for random selection.

When considered over one generation the criterion of selection can be maximized by differential use of selected animals, however in this study a fixed number of candidates were selected and used equally in each generation. This is because it is not yet possible to predict the change in information flow over several generations, and the assumption made in this study ensures information remains in a

predictable equilibrium. For example, if differential use were allowed then the best individual would be used the most and worst the least (Toro and Nieto, 1984). If error variance were penalized, then in subsequent generations the offspring of the worst would be doubly penalized in the decision processes considered through (i) inheritance, and (ii) fewer relatives with consequently greater error variance. Models to cope with either of these consequences require an understanding of line proliferation through generations of selection or equivalently an understanding of inbreeding under selection.

Application in practice

If a linear or quadratic loss function were to be incorporated into selection decision it would be necessary to consider the value of the parameter α that would be assumed. For linear loss the interpretation of the selection criterion as a percentile can be readily viewed in terms of risk. If the 25%-ile is used then an individual or group has a probability of 0·75 of achieving the gain indicated by the selection criterion (i.e. the strategy has low risk) whereas if the 75%-ile is used the indicated gain or more is achieved only with probability 0·25 (a strategy of high risk, but gains when achieved are greater). From these considerations a value of α to suit the desired risk can be derived easily. For quadratic loss the value of α is not so easily defined; it would seem necessary to choose α on the basis of the relationship between genetic gain and the variation of response obtained from deterministic (as in this study) or simulation models. Although the interpretation of quadratic loss is less clear than the linear loss its benefit comes from the observation that the problem can be formulated in terms of quadratic programming for which computational algorithms exist. Nevertheless, the computational power required is still a very severe limitation on its use. The approach adopted by Schneeberger *et al.* (1982) was equivalent to assuming a quadratic loss function for a single generation of selection.

The decisions that need to be taken in practice that affect the variance of response are firstly the number of parents (Hill, 1971), followed by choices among individuals of differing accuracies, and choices among related and unrelated individuals. Increasing the number of parents, although it is likely to reduce the expected genetic gain, will offset extra risk from using parents of low accuracy, and this has been used in the marketing of young unproven bulls (Genus, 1991). For a given number of parents, a comparison of the results examples 3 and 4 shows that accounting for differing accuracies alone can have a substantial impact on the variance of responses, and that over and above this the covariances between relatives have only a small impact. Furthermore, decisions involving the selection of relatives need to be considered in the wider context of inbreeding. Neglecting the covariances between relatives considerably simplifies the computational problems to a comparison involving different numbers of individuals on their accuracies alone. This could then be applied in the context of profit functions involving either estimation errors or thresholds.

In conclusion, whilst previous studies (e.g. Schneeberger *et al.*, 1982; Van Raden, Freeman and Rothschild, 1984) have considered the adjustment of selection criteria for accuracy, this study has (i) set up a framework for considering the rationality of criteria, (ii) considered a range of alternative criteria, and (iii) described the effect of these criteria on both genetic gain and variance of gain, the latter being one aspect of risk in breeding schemes. When applied to dairy breeding schemes, where accuracies of evaluation of bulls may vary widely, it would seem straightforward to derive percentiles from the posterior distribution of the breeding value and for catalogues to present low risk, expected and high risk estimates of breeding value for each bull. The levels of risk chosen should ideally be common to all catalogues and producers may then select according to their desired level of risk. In such a way producers can begin to make appropriate choices between bulls differing in their accuracy of evaluation.

References

Ferguson, T. S. 1967. *Mathematical statistics, a decision theoretic approach*, pp. 44-45. Academic Press, London.

Genus. 1991. Sire catalogue.

Goffinet, B. and Elsen, J. M. 1984. Optimal selection criteria: some general results. *Genetique, Selection, Evolution* **16:** 307-318.

Hill, W. G. 1971. Design and efficiency of selection experiments. *Biometrics* **1972:** 293-311.

Meuwissen, T. H. E. 1990. The effect of the size of MOET nucleus dairy cattle breeding plans on the genetic gain and its variance. *Proceedings of the fourth world congress on genetics applied to livestock production, Edinburgh, vol. XIV*, pp. 271-274.

Meuwissen, T. H. E. 1991. Expectation and variance of genetic gain in open and closed nucleus and progeny testing schemes. *Animal Production* **53:** 133-141.

Nicholas, F. W. 1989. Incorporation of new reproductive technology in genetic improvement programs. In *Evaluation and animal breeding* (ed. W. G. Hill and T. F. C. Mackay). CAB International, Wallingford.

Nicholas, F. W. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

Schneeberger, M., Freeman, A. E. and Boehlje, M. D. 1982. Application of portfolio theory to dairy sire selection. *Journal of Dairy Science* **65:** 404-409.

Toro, M. A. and Nieto, B. M. 1984. A simple method for increasing the response to artificial selection. *Genetical Research* **44:** 347-349.

Woolliams, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49:** 1-14.

Woolliams, J. A. 1990. Strategies to maximise selection progress in dairy cattle. *Proceedings of the fourth world congress on genetic applied to livestock production, Edinburgh, vol. XIV*, pp. 15-24.

Woolliams, J. A. 1991. A successor to MOET? *British Cattle Breeders Club Digest no. 45*, pp. 8-13.

Van Raden, P. M., Freeman, A. E. and Rothschild, M. F. 1984. Maximising genetic gain under multiple-stage selection. *Journal of Dairy Science* **67:** 1761-1766.

(*Received* 1 *July* 1992—*Accepted* 21 *October* 1992)

Paper 18

Maximizing Genetic Response in Breeding Schemes of Dairy Cattle with Constraints on Variance of Response

T.H.E. MEUWISSEN
Research Institute for Animal Production 'Schoonoord'
PO Box 501
3700 AM Zeist, The Netherlands

J. A. WOOLLIAMS
Roslin Institute (Edinburgh)
Roslin
Midlothian EH25 9PS, Scotland

ABSTRACT

Predicted genetic progress in dairy cattle breeding schemes was maximized with the variances of selection responses constrained, i.e., restricted effective population sizes. This restriction would also lead to schemes with acceptable rates of inbreeding (<.5%/yr).

If the required coefficient of variation of the annual selection response was reduced from .32 to .16, numbers of animals selected, openness of schemes, and generation intervals increased. When elite cows produced 8 offspring annually, this reduction of the coefficient of variation tended toward a conventional progeny-testing scheme. If the number of donor cows was optimized, responses increased ≤2%, and the breeding schemes became virtually closed. Variances of responses were reduced by selecting fewer, but proven, bulls, as is done in hybrid multiple ovulation and embryo transfer schemes, which select progeny-tested bulls and young elite cows.

In spite of the constrained coefficients of variation, maximized genetic gains were high and were only reduced from .300 to .293 genetic standard deviations per year, when coefficients of variation were reduced from .32 to .16. Adoption of breeding schemes with low coefficients of variation is recommended, because responses are high and coefficients of variation are sensitive to accidental changes in the breeding structure.

Received June 21, 1993.
Accepted February 16, 1994.

(Key words: breeding schemes, restricted maximization of selection response, variance of selection response, simulated annealing)

Abbreviation key: MOET = multiple ovulation and embryo transfer.

INTRODUCTION

Optimization of breeding schemes of dairy cattle has generally been for increased genetic gain with only an implicit restriction on the risk of the breeding scheme (2, 3, 11, 13, 16). Components of this risk are variance of the response and inbreeding, which both increase as the numbers of sires selected decrease. For this reason, the restriction most commonly and most easily applied has been to fix the number of sires selected annually prior to beginning the optimization. However, annual rates of inbreeding and variance of response are also affected by generation intervals, accuracy of selection, coselection of sibs, and the number of dams selected. The introduction of **MOET** (multiple ovulation and embryo transfer) (16) and BLUP (7) has led to the development of a variety of schemes that differ greatly in some or all of these aspects. Optimization of breeding schemes that have an explicit constraint on variance of response or inbreeding and that do not fix a priori the number of sires selected is therefore very much needed. Woolliams (20), Woolliams and Wilmut (23), and Quinton et al. (18) constrained rates of inbreeding when comparing mating designs, new technologies, and BLUP with phenotypic selection, respectively.

The debate on the selection of young versus older animals or, equivalently, MOET versus

progeny-testing schemes is still open if schemes are compared at equal rates of inbreeding or by variances of responses. Nicholas and Smith (16) found increases in genetic gain of ≤50% when they compared MOET schemes with conventional progeny-testing schemes. This increase in genetic gain was mainly due to the markedly reduced generation intervals. However, with selection of a constant number of animals per year, inbreeding is proportional to the reciprocal of the squared generation interval because, when generation interval is smaller, fewer animals per generation are selected, and more selection cycles are undertaken per unit of time. Other compounding effects occur from the selection of younger bulls because selection is often based on indices containing only family information, which increases the probability of coselection of siblings (24). To compare MOET schemes and conventional progeny-testing schemes at equal rates of inbreeding or variances of responses, more animals should be selected in the MOET schemes, which decreases their rates of genetic gain.

Nicholas (15) argued that obtaining acceptable variation of responses forms a more stringent restriction on effective population sizes than does obtaining acceptable rates of inbreeding. Depending on the degree of risk aversion, coefficients of variation of the response of 10 yr of selection of .05 to .10 and annual rates of inbreeding of .005 were considered to be acceptable. For an adult MOET scheme, Nicholas (15) found that the former constraint led to effective population sizes of 100 to 400 and the latter to 27 animals per generation. Hence, schemes that satisfy constraints on variance of response implicitly lead to acceptable rates of inbreeding, but the converse may not be true.

The aim of this study was to maximize selection response in breeding schemes of dairy cattle with an explicit constraint on its coefficient of variation. Optimization is for the number of animals selected, the ages at which selection takes place, and the openness of the nucleus. Open and closed nucleus schemes, with and without progeny testing of young bulls, were considered. In this optimization, open nucleus schemes with progeny testing of young bulls and selection of (few) old animals represent the conventional progeny-testing schemes.

MATERIALS AND METHODS

Breeding Schemes

The parameters of the breeding schemes are given in Table 1. Selection is for lactation production, an aggregate trait that might in-

TABLE 1. Parameters of the breeding schemes.

Size of nucleus, no. neonates/yr	256 ♂ + 256 ♀
Size of base (commercial cow population)	1,000,000
Test daughters per progeny-tested bull, no.	100
Progeny per elite dam, no./yr	8 or optimized[1]
Age classes available for selection of sires and dams, yr	2–10
No. selected in path	
Sires to breed nucleus replacements, no.	optimized
Sires to breed base replacements	same sires as those to breed nucleus replacements
Dams to breed nucleus replacements, no.	64 or optimized
Dams to breed base replacements	random selection
Parameters of milk production	
Heritability	.25
Repeatability	.40
Genetic correlation between lactation records	1
Involuntary culling	
Bulls, %	0[2]
Cows, %/yr	30

[1]Optimized indicates that the number of offspring per elite cow is not restricted as may be the case with in-vitro maturation and fertilization techniques (5).

[2]Semen supplies take over the role of a deceased bull.

clude milk, fat, and protein yield. The categories of breeding schemes considered were category 1, closed nucleus without progeny testing and with 8 progeny per elite dam; category 2, as category 1 but with progeny testing; category 3, as category 2 but with an open nucleus; categories 4, 5, and 6 repeated categories 1, 2, and 3 with the restriction on the number of offspring per elite dam removed completely and allowing an unlimited number of offspring per elite dam. In this context, "unlimited" means that the nucleus replacements can be produced by very few females, which becomes feasible with in vitro maturation and in vitro fertilization or other techniques in the future (5). In these latter categories, the optimized number of elite dams selected might be smaller than the optimized number of sires, which leads to the mating of several bulls to the same cow and a maternal half-sib family structure. If the number of sires is larger than the number of cows, several cows are mated to the same bull, and a paternal half-sib family structure results.

The same sires were selected to breed nucleus and base replacements to reduce the size of the optimization problem. If the number of sires selected is very low, selection of a few more sires to breed the 1,000,000 base replacements will probably decrease rates of genetic gain only marginally. Involuntary culling of sires was neglected, because stored semen can take over the role of a deceased bull. In the path of males to breed base replacement, this assumption may be somewhat optimistic, but, in practice, involuntary culling rates of bulls are low. The selection response in the path of dams to breed base replacement was assumed negligible because of the high proportion of cows selected in this path and because selection is not only for breeding value estimates of lactation production in commercial herds.

The Optimization Algorithm: Simulated Annealing

The method of simulated annealing is particularly useful for large problems of combinatorial optimization, which consist of finding, from among a set of alternatives, one that maximizes the value of an objective function (17). The present problem is combinatorial because the optimal combination of numbers of animals selected from each age class has to be found. Selection of 0, 1, 2, . . ., or 512 bulls from each age class is possible. Nine age classes are eligible for selection (age classes 2, . . ., 10); hence, the total number of possible combinations of predefined numbers of animals selected per age class is $513^9 - 1$. If the number of dams to select from each age class is also to be optimized, the number of combinations is approximately squared in a closed scheme and more than cubed in an open scheme. Hence, the combinatorial problem is very large indeed. This problem is usually solved by applying the same truncation point across age classes for the selection of the predefined number of animals, such as in the work of Meuwissen (11). This procedure may lead to large coefficients of variation and requires a predefined total number of animals selected.

The simulated annealing algorithm has an analogy in physics that is described here because it aids understanding of the process and introduces some of the terminology that has become associated with the algorithm. Annealing refers to the slow cooling of metal in order to harden it (17). During this process, the atoms find their optimal positions to achieve a state of minimum energy. At the beginning, individual atoms move in all directions, which may either increase or decrease the energy, but, as temperature decreases, directions that increase energy become less probable. In this way, the order with minimum energy is found for billions of atoms.

This optimization process, which occurs in nature, is simulated by the present algorithm. An objective function, Ω, which is to be minimized, is defined, and an initial solution is generated. In the case of cooling of metal, the objective function is the energy level, and a solution refers to the position of each atom. In the present case, a solution represents the number selected from each age class. At each iteration, a small random modification to the current solution is suggested (a change in the position of a single atom). If the modification decreases the object function, it is accepted and replaces the current solution. Otherwise, the modification is accepted with a probability that depends not only upon how much poorer the solution is but also upon an imaginary temperature; the poorer the modification is,

and the cooler the temperature, the less likely the modification is to be accepted. The probability that a change is accepted is from the Boltzmann probability distribution of a system (e.g., metal) being at a certain level of energy (17) and is

$$P_{acc} = \exp(-\Delta\Omega/T)$$

for $\Delta\Omega > 0$, and

$$P_{acc} = 1$$

for $\Delta\Omega \leq 0$,

where P_{acc} = probability of accepting the modification, $\Delta\Omega$ = change of the object function value, and T = imaginary temperature, which was initially set to .03. This initial value of T is of the same order of magnitude as the changes in genetic gain because of modifications, which leads to acceptance of most initially suggested modifications. As in results of Press et al. (17), the imaginary temperature was multiplied by .9 each time 20 or 40 modifications (10% of the number of age classes) were accepted for closed and open schemes, respectively. The optimal solution is found when the solution is not changed during 200 or 400 suggested modifications, respectively. If the solution does not change further, the current solution is very unlikely to be improved upon, and the temperature has reached a sufficiently low value.

In the present case, we want to maximize expected genetic gain, E(ΔG), which is identical to the problem of minimizing -E(ΔG). To apply the algorithm, we take a solution to be the number selected from each age class. Modifications to solutions were chosen at random from the following three: 1) select one additional animal from a selection path and age class chosen at random, 2) decrease the number of animals selected by one in a selection path and age class chosen at random, and 3) perform modifications 1 and 2 simultaneously for two age classes and one selection path chosen at random. Modification 3 changes the distribution of the animals selected for a selection path and keeps the total number of animals selected the same. This modification is required because an improvement in the objective function might result even though the single component changes, one of type 1 and the other of type 2, may both decrease the objective function.

The objective function is chosen such that, in the optimal scheme, genetic gain is maximized, and the coefficient of variation of the response does not exceed CV_c (the critical coefficient of variation):

$$\Omega = -E(\Delta G),$$

if $V(\Delta G) \leq CV_c^2\ E^2(\Delta G)$, and

$$\Omega = -E(\Delta G) + k(V(\Delta G) - CV_c^2\ E^2(\Delta G)),$$

if $V(\Delta G) > CV_c^2\ E^2(\Delta G)$, where $E(\Delta G)$ and $V(\Delta G)$ = expectation and variance of genetic gain, respectively, and k = large positive constant. In the present study, k = 10,000 was sufficiently large to give only coefficients of variation of optimized schemes $\leq CV_c$.

The simulated annealing algorithm is not guaranteed to find the global optimum, but the algorithm does not converge rapidly to an unfavorable local optimum as many other optimization techniques do (17). Generally, the algorithm finds a solution that is close to optimum, and significantly better solutions are unlikely to be found.

The Model

The basic model used was that of Meuwissen (13). In that model, the optimization of generation intervals maximized genetic gain, and improvement at each stage depended on the genetic gain achieved, which made an iterative algorithm necessary for calculating genetic gains. The optimization of the generation interval part of the model was omitted here, because generation intervals are defined by the annealing algorithm, which made possible faster calculation of $E(\Delta G)$ and $V(\Delta G)$. Also, fast calculation of $E(\Delta G)$ and $V(\Delta G)$ was necessary because the annealing algorithm required many evaluations of $E(\Delta G)$ and $V(\Delta G)$.

Let I_{xi} denote the selection index for animal category x and age class i, where each age class contains 1 yr, and categories of animals are SN, DN(N), DN(B), SB, and DB, which denote, respectively, sires to breed nucleus replacements, dams to breed nucleus replacements selected from the nucleus, DN selected from the base, sires to breed base replace-

ments, and dams to breed base replacements. Age class i is defined such that a selected animal in age class i will have offspring when it is i yr old (selected animals in age class i have a generation interval of i yr). The selection indices combine records available on the individual, its progeny, half- and full siblings, its parents, their half- and full sib, and its grandparents. Separate indices were calculated for males and females within each age class, within nucleus and base. Within sex, age, and tier classes, the same amount of information was assumed to be available for each individual. The extent of the information that was included enables the indices to approximate BLUP closely.

Mean values of selection indices of the selected animals were from

$$\bar{I}_x = \Sigma_i c_{xi} \bar{i}_{xi} \sigma_{Ixi}(\infty),$$

where summation is over all age classes i of animal category x; c_{xi} = contribution of age class i to the animal category x, which is from the annealing algorithm; $\bar{i}_{xi}$ = selection intensity, which is corrected for finite population size and correlations between breeding value estimates of half- and full sibs (12); $\sigma^2_{Ixi}(t)$ = variance of the selection index of animal category x and age class i in yr t. $\sigma^2_{Ixi}(\infty)$ = equilibrium variance, which is reduced because of selection (1). Equilibrium variances are obtained from $\sigma^2_{Ixi}(\infty) = \sigma^2_a(\infty) - PEV_{xi}$, where $\sigma^2_a(\infty)$ = equilibrium additive genetic variance and PEV_{xi} = prediction error variances of estimated breeding values. Values of $\sigma^2_a(\infty)$ are obtained by a few iterations on formulas of Bulmer (1) until successive values converge. Because prediction error variances (PEV) are not affected by selection (6), $PEV_{xi} = \sigma^2_a(0) - \sigma^2_{Ixi}(0)$, where 0 denotes the unselected founder population.

With generation intervals defined by the annealing algorithm, genetic gains of open nucleus schemes are predicted by (11):

$$E(\Delta G) = \Sigma\bar{I}/\Sigma\bar{L} =$$

$$\frac{\bar{I}_{SN} + f_{NN}\bar{I}_{DN(N)} + (1 - f_{NN})\bar{I}_{DN(B)} + \bar{I}_{SB} + \bar{I}_{DB}}{\bar{L}_{SN} + f_{NN}\bar{L}_{DN(N)} + (1 - f_{NN})\bar{L}_{DN(B)} + \bar{L}_{SB} + \bar{L}_{DB}} \quad [1]$$

where $E(\Delta G)$ = expected steady state selection response; $\bar{L}_x$ = mean generation interval of selected animals of animal category x, respectively; and f_{NN} is the fraction of DN selected from the nucleus. If $f_{NN} = 1$, this equation reduces to the well-known formula for closed nucleus schemes: $E(\Delta G) = [\bar{I}_{SN} + \bar{I}_{DN(N)}]/[\bar{L}_{SN} + \bar{L}_{DN(N)}]$. And, with $f_{NN} = 0$, the formula for a progeny-testing scheme is obtained: $E(\Delta G) = [\bar{I}_{SN} + \bar{I}_{DN(B)} + \bar{I}_{SB} + \bar{I}_{DB}]/[\bar{L}_{SN} + \bar{L}_{DN(B)} + \bar{L}_{SB} + \bar{L}_{DB}]$ (19).

The approach of Johnson (10) was followed to derive a formula for the variance of the selection response $V(\Delta G)$. Let $\mathbf{z}_t$ denote a vector of length 20 with mean genetic merits of the first 10 age classes (age classes >10 yr are neglected) in the nucleus and the base at year t, and let $\mathbf{z}_t(1)$ and $\mathbf{z}_t(11)$ denote the first nucleus and base age class, respectively. Further, let $\mathbf{s}_t$ denote a vector with genetic selection differentials, which can be obtained by weighting the $\bar{I}_x$ values over categories of selected animals. The exact form is not important here. The vector $\mathbf{z}_{t+1}$ can be expressed in parameters of year t:

$$\mathbf{z}_{t+1} = \mathbf{P}\mathbf{z}_t + \mathbf{s}_t + \mathbf{e}_t, \text{ for } t \geq 1, \quad [2]$$

where $\mathbf{P}$ = a matrix that describes the flow of genes from year t to t + 1 (8), and $\mathbf{e}_t$ = deviations from the expected genetic level because of sampling. For an open nucleus scheme, the gene flow matrix $\mathbf{P}' = [\mathbf{p}_N\ \mathbf{u}_1\ \mathbf{u}_2 \ldots \mathbf{u}_9\ \mathbf{p}_B\ \mathbf{u}_{11}\ \mathbf{u}_{12} \ldots \mathbf{u}_{19}]$, where $\mathbf{p}_N$ = vector of contributions all nucleus and base age classes to nucleus ($p_N(i) = \frac{1}{2}(c_{SNi} + f_{NN}c_{DN(N)i})$, for $1 \leq i \leq 10$; $p_N(i) = \frac{1}{2}(1 - f_{NN})c_{DN(B)i}$, for $11 \leq i \leq 20$); $\mathbf{p}_B$ = vector of contributions all nucleus and base age classes to base ($p_B(i) = \frac{1}{2}c_{SBi}$, for $1 \leq i \leq 10$; $p_B(i) = \frac{1}{2}c_{DBi}$, for $11 \leq i \leq 20$); and $\mathbf{u}_i$ = vector with all zeros and a 1 at position i.

The variance-covariance matrix of $\mathbf{e}_t$ is denoted by $\mathbf{V}_e$ and has only nonzero elements at positions (1,1), (11,1), (1,11), and (11,11). Elements $V_e(i,i)$, with $2 \leq i \leq 10$ or $12 \leq i \leq 20$, are assumed to be zero because the mean merit of age class i - 1 in year t - 1 is approximately equal to that of class i in yr t, if culling is not correlated with lactation yield. Similarly, $V_e(i,j) = 0$, where $2 \leq i \leq 10$ or $12 \leq i \leq 20$ and $1 \leq j \leq 20$. From Meuwissen (13),

$$V_e(1,1) = \frac{Var(\Sigma_i TBV_{SNi})}{4n_{SN}^2} + \frac{Var(\Sigma_i TBV_{DNi})}{4n_{DN}^2} + \frac{Cov(\Sigma_i TBV_{SNi}\Sigma_i TBV_{DNi}}{2n_{SN}n_{DN}}$$

where TBV_{xi} = true breeding value of animal i selected for category x; and n_x = number of animals selected for category x. The value of $V_e(11,11)$ is calculated similarly with SN and DN replaced by SB and DB. Further,

$$V_e(1,11) = \frac{Cov(\Sigma_i TBV_{SNi};\Sigma_i TBV_{SBi})}{4n_{SN}n_{SB}} + \frac{Cov(\Sigma_i TBV_{SNi};\Sigma_i TBV_{DBi})}{4n_{SN}n_{DB}} + \frac{Cov(\Sigma_i TBV_{DNi};\Sigma_i TBV_{SBi})}{4n_{DN}n_{SB}} + \frac{Cov(\Sigma_i TBV_{DNi};\Sigma_i TBV_{DBi})}{4n_{DN}n_{DB}}$$

Evaluation of, for instance, $Var(\Sigma_i TBV_{SNi})$ requires calculation of

$$Cov(TBV_{SNi};TBV'_{SNj}) =$$

$$Cov(I_{SNi};I_{SNj}) + PEC_{SN;i,j}, \qquad [3]$$

where $PEC_{SN;i,j}$ = prediction error covariance of i and j. The terms in Equation [3] depend on the family relationship of the selected sires i and j. Only full and half-sib relationships are considered. Selection of animals within an age class was based on expected order statistics of their selection index values (13). If this selection led to the coselection of full or half-sibs, the $Cov(I_{SNi};I_{SNj})$ and $PEC_{SN;i,j}$ of full or half-sibs, respectively, were used in Equation [3]. The other terms, $V_e(1,1)$, $V_e(1,11)$, and $V_e(11,11)$, were calculated similarly.

Because e_t in Equation [2] represents the variable part of selection response, $Var(z_{t+1}) = Var(\Sigma_{i=0}^{t} P^i e_i)$. Hence, $Var(z_{t+1} - z_t) = Var(P^t e_t) = P^t V_e P'^t$. Hill (8) shows that $\lim_{t\to\infty} P^t = qv'/(2\Sigma\bar{L})'$, where $\Sigma\bar{L}$ is as defined by Equation [1], v is the left eigenvector of P associated with its largest eigenvalue, and q is a vector of ones. The variance of the steady-state selection response is $\lim_{t\to\infty} Var(z_{t+1} - z_t) = \lim_{t\to\infty} P^t V_e P'^t = \frac{1}{4}(\Sigma\bar{L})^{-2}\, v'V_e v^* qq'$; hence, the variances of all elements of z_t increases at the same rate, which is the variance of the selection response. Prediction of the nonzero elements (1,1), (1,11), (11,1), and (11,11), of V_e has been described; hence, only the elements v(1) and v(11) are still required to calculate $v'V_e v$. Following a method of Hill (8), v(1) = 1, and v(11) = $(1 - f_{NN})$. Consequently, the steady-state variance of the selection response is

$$V(\Delta G) = \frac{1}{4}(\Sigma\bar{L})^{-2}\,(V_e(1,1) + 2(1 - f_{NN})V_e(1,11) + (1 - f_{NN})^2 V_e(11,11)). \qquad [4]$$

Predicted variances of responses from this formula were virtually identical to the converged steady-state variances obtained from the model of Meuwissen (13). Computing times were much reduced if Equation [4] was used, which was desirable because many evaluations of V(ΔG) were required.

The correction of selection differentials for correlations between expected breeding values of full and half-sibs by the method of Meuwissen (12) requires a hierarchical breeding structure; i.e., each male is mated to several females, leading to a paternal half-sib family structure, or each female is mated to several males, leading to a maternal half-sib family structure. Input parameters for this method are the fraction selected, the number of half-sib families (n_{HS}), the number of full-sib families per half-sib family (n_{FS}), the number of males (or females) within a full-sib family (n_w), and intraclass correlations between full and half-sibs. With a paternal family structure, n_{HS} = the number of sires selected, n_{FS} = the number dams divided by the number of sires, and n_w = the total number of males (or females) divided by the number of dams. If n_{FS} and n_w were noninteger, n_{FS} and n_w were rounded to their nearest integer. Rounding only affected the family structure; fractions selected were not affected, hence, approximately accounting for the effect of family structure on intensities of selection.

RESULTS

We did not attempt to optimize the parameters of the annealing algorithm, e.g., start temperature, number of modifications within each temperature step. The number of evaluations of breeding schemes was large: 2000 to 4000 for closed nucleus and 7000 to 10,000 for open nucleus schemes. For a few schemes, the optimization was performed twice, and numbers of animals selected differed by ≤2 (results not shown). Solutions were not identical because suggested modifications were sampled at random. Genetic gains of the alternative schemes differed by <.1%.

The coefficient of variation of the annual selection response was restricted to .32 or .16; thus, the coefficient of variation of the response of 10 yr of selection was .1 or .05, respectively ($CV(t) = CV(1)/\sqrt{t}$, where CV(t) is coefficient of variation of responses of t year of selection). These figures were chosen because they had also been considered by Nicholas (15). The coefficient of variation of the optimized breeding schemes was, in all but one scheme, very close to its constraints.

Table 2 shows the results of the optimization for the schemes for which the number of elite cows to breed nucleus replacements was fixed at 64 with 8 progeny per elite cow. If the coefficient of variation was reduced from .32 to .16, expected genetic gains decreased by 11 and 5% for closed nucleus schemes without and with progeny tests, respectively. In the scheme without progeny testing, the reduction of the coefficient of variation was mainly achieved by increasing the number of sires selected; in the scheme with progeny testing, generation intervals were also increased. Paradoxically, this increase resulted in a smaller reduction in genetic gain in the progeny-testing schemes because the necessary changes in the number of bulls selected was less dramatic.

In the open nucleus schemes, the reduction of the coefficient of variation decreased response only 3% (Table 2). These schemes used the same means as the closed scheme with progeny testing but were more open, i.e., smaller f_{NN}. Variation of response decreased with decreasing f_{NN}, because ΣL in Equation [3] became substantially larger (see Equation [1]). By decreasing f_{NN}, the genetic gain in open nucleus schemes was reduced less than in closed schemes (Table 2). Table 3 shows the number of animals selected annually from each age class and path in the open nucleus scheme. The exact number of animals selected from each age class is of less interest than the qualitative changes in the age structure of selected animals when the coefficient of variation constraint is changed. The number of selected bulls that were progeny-tested (age

TABLE 2. The maximized expected response and corresponding structure of breeding schemes with 64 elite cows to breed nucleus replacements, for different required coefficients of variation (CV_c).

CV_c	$E(\Delta G)$[1]	$CV(\Delta G)$	N_{SN}[2]	L_{SN}[3]	L_{DN}[3]	f_{NN}[4]
	Closed nucleus without progeny testing of young bulls					
.32	.294	.316	12	2.7	2.3	1
.16	.262	.160	44	3.7	2.7	1
	Closed nucleus with progeny testing of young bulls					
.32	.297	.287	13	2.9	2.2	1
.16	.281	.160	20	4.3	2.4	1
	Open nucleus with progeny testing of young bulls					
.32	.300	.315	6	3.0	2.1	.67
.16	.293	.160	19	3.5	2.2	.47

[1]Expected genetic gain in σ_{a0}-units per year.

[2]Number of sires to breed nucleus replacements selected. In open nucleus schemes, N_{SN} equals also the number males to breed base replacements.

[3]Generation intervals of sires to breed nucleus replacements (L_{SN}) and dams to breed nucleus replacements (L_{DN}) in year.

[4]The proportion of females to breed nucleus replacements that are selected from the nucleus.

TABLE 3. The number of animals selected per age class in an open nucleus scheme with progeny testing and 64 elite dams, when expected selection response is maximized for different required coefficients of variation (CV_c).[1]

	SN[3]		DN(N)[3]		DN(B)[3]	
Age class[2]	CV_c = .32	CV_c = .16	CV_c = .32	CV_c = .16	CV_c = .32	CV_c = .16
2	4	8	36	24	21	31
3	0	1	5	4	0	0
4	1	7	2	2	0	3
5	0	0	0	0	0	0
6	1	2	0	0	0	0
7	0	1	0	0	0	0

[1]The maximum number of animals selected per age class is 256 for bulls and follows from the number of neonates and involuntary culling rates for cows.

[2]Age is at birth of offspring of the selected animals.

[3]SN = Sires to breed nucleus replacements; DN(N) and DN(B) = dams to breed nucleus replacements selected from the nucleus and base, respectively.

classes 6 and 7) increased when the coefficient of variation was reduced; however, even with coefficient of variation constraint of .16, the majority of the bulls were young bulls (age class 2 and 3) and sib-tested bulls (classes 4 and 5). The elite cows selected were mainly young heifers. The reduction of the coefficient of variation hardly increased the generation interval for elite cows but increased substantially the number of elite cows selected from the commercial cow population.

When the total number of elite cows was allowed to vary and a paternal half-sib hierarchical structure was imposed, genetic gains increased only 2%. Table 4 shows the results for the three types of breeding schemes under consideration. The increased reproductive rate of elite cows led to nucleus schemes that were virtually completely closed.

When the coefficient of variation constraint was reduced, schemes with progeny testing selected fewer males per year but increased the

TABLE 4. The maximized expected response and corresponding structure of breeding schemes when the number of elite cows required is allowed to vary, for different required coefficients of variation (CV_c).

CV_c	E(ΔG)[1]	CV(ΔG)	N_{SN}[2]	N_{DN}[2]	L_{SN}[3]	L_{DN}[3]	f_{NN}[4]
	Closed nucleus without progeny testing of young bulls						
.32	.297	.320	17	34	3.1	2.4	1
.16	.266	.160	37	75	3.4	2.6	1
	Closed nucleus with progeny testing of young bulls						
.32	.304	.319	16	32	3.3	2.2	1
.16	.288	.160	4	30	6.3	2.3	1
	Open nucleus with progeny testing of young bulls						
.32	.304	.319	16	32	3.3	2.2	1
.16	.291	.159	4	30	6.3	2.1	.97

[1]Expected genetic gain in σ_{a0} units per year.

[2]N_{SN} and N_{DN} = Number of sires and dams to breed nucleus replacements, respectively. In open nucleus schemes, N_{SN} equals also the number sires to breed base replacements. To obtain a paternal half-sib family structure, the condition $N_{SN} \leq N_{DN}$ was enforced.

[3]Generation intervals of sires to breed nucleus replacements (L_{SN}) and dams to breed nucleus replacements (L_{DN}) in year.

[4]The proportion of females to breed nucleus replacements that are selected from the nucleus.

TABLE 5. The number of animals selected per age class in an open nucleus scheme with progeny testing and variable numbers of both bulls and cows, when expected selection response is maximized for different required coefficients of variation (CV_c).[1]

	SN[3]		DN(N)[3]		DN(B)[3]	
Age class[2]	CV_c = .32	CV_c = .16	CV_c = .32	CV_c = .16	CV_c = .32	CV_c = .16
2	8	0	28	27	0	1
3	0	0	3	1	0	0
4	6	0	1	1	0	0
5	0	0	0	0	0	0
6	2	3	0	0	0	0
7	0	1	0	0	0	0

[1]The maximum number of animals selected per age class is 256 for bulls and follows from the number of neonates and involuntary culling rates for cows (see Table 1).

[2]Age is at birth of offspring of the selected animals.

[3]SN = Sires to breed nucleus replacements; DN(N) and DN(B) = dams to breed nucleus replacements selected from the nucleus and base, respectively.

number of progeny-tested bulls (Tables 4 and 5). Without progeny testing, more sires and dams were selected when coefficient of variation constraint was reduced and the genetic gain decreased by 10% compared with the 5% decrease in progress for schemes with progeny testing. Thus, for an open nucleus scheme, a strong constraint on the coefficient of variation and a high reproductive rate of elite cows led to the use of progeny-tested bulls in a virtually closed nucleus scheme. This scheme resembles very closely the hybrid MOET schemes proposed by Colleau (2).

In schemes with unlimited reproductive rates of elite cows, the number of elite cows selected might be smaller than the number of males, which led to maternal half-sib family structures. Evaluations of these schemes were possible by the presented model as long as the mating design was hierarchical and the nucleus was closed. The latter restriction was due to the design of the computer program but was not a severe limitation in practice because schemes with selection of few dams were closed anyway (Table 4).

Table 6 shows the results for the schemes with maternal half-sib families. In the absence of progeny testing, genetic gains were up to 3% higher than that of the schemes with paternal half-sib families mainly because dams in age classes 4 and 5 have higher accuracy of selection than bulls of the same age, which is then combined with a higher intensity of selection of dams in the maternal schemes. Greater rates of gain for maternal than for paternal schemes were also found by De Boer and van Arendonk (3) for adult MOET schemes. The number of animals selected and generation intervals of the maternal schemes are very similar to those of their corresponding paternal schemes; the dams replaced the sires and vice versa. With progeny testing and coefficient of variation constraint of .16, the superior paternal scheme used only a few proven bulls. This scheme resulted in a 4% higher genetic gain of the paternal than the maternal scheme.

DISCUSSION

The number of animals selected and the generation intervals were optimized in breeding schemes that constrained variances of selection responses. This optimization should result in schemes with both acceptable variances of selection responses and acceptable rates of inbreeding. Maximized genetic gains were high, about .3 and .28 to .29 genetic standard deviation units per year, and coefficients of variation were .32 and .16, respectively. These rates of gain were as high as the highest ranking scheme of Meuwissen (13), which had a genetic gain of .299 σ_{a0} units/yr and a coefficient of variation of .66. Hence, the optimization resulted in remarkably high genetic gains when coefficients of variation

TABLE 6. The maximized expected response and corresponding structure of breeding schemes when the number of elite cows required is allowed to vary, for different required coefficients of variation (CV_c).

CV_c	$E(\Delta G)^1$	$CV(\Delta G)$	N_{SN}^1	N_{DN}^1	L_{SN}^3	L_{DN}^3
	Closed nucleus without progeny testing of young bulls					
.32	.305	.319	36	17	2.2	2.9
.16	.271	.160	76	28	2.6	3.5
	Closed nucleus with progeny testing of young bulls					
.32	.307	.320	36	13	2.3	3.1
.16	.276	.160	74	30	2.7	3.3

[1]Expected genetic gain in σ_{a0}-units per year.

[2]N_{SN} and N_{DN} = Number of sires and dams to breed nucleus replacements, respectively. In open nucleus schemes, N_{SN} equals also the number sires to breed base replacements. In order to obtain a maternal half-sib family structure, the condition $N_{DN} \leq N_{SN}$ was imposed.

[3]Generation intervals of sires to breed nucleus replacements (L_{SN}) and dams to breed nucleus replacements (L_{DN}) in year.

were constrained, mainly because the number of animals selected per path was not optimized by Meuwissen (13). Also, Meuwissen's (13) main interest was in MOET schemes with short generation intervals; hence, young nucleus bulls were not progeny tested.

Prediction of genetic gains and its variances ignored inbreeding. At a fixed coefficient of variation of the selection response, the rates of inbreeding would be similar because a close relationship exists between variance of response and inbreeding (9) and rates of gain are similar, and thus the schemes would be affected in the same way. Because rates of genetic gain for schemes with coefficients of variation of .16 and .32 were similar, we might assume that schemes with a coefficient of variation of .16 have reduced rates of inbreeding and that, over the long term, genetic gains would be more favorable than those of schemes with a coefficient of variation of .32.

Woolliams and Meuwissen (21) decreased the variance of the selection response by selecting for estimated breeding values minus a factor k times their prediction error variances, thus penalizing for prediction errors. Variance of response is not only due to prediction errors on breeding value estimates but is also due to Mendelian sampling of genes. The latter is reduced by reducing the number of samplings per unit of time (i.e., increasing generation intervals, or by selecting more animals, which reduces sampling effects). The method of Woolliams and Meuwissen (21) does not aim at a predefined variance of the response, and it did not optimize the number of animals selected, although it could be easily implemented in practice. The present method could be implemented by using the approximate numbers of animals selected given in Tables 3 or 5, but this method is not optimal because of sampling effects. A sequential optimization rule for controlling risk, such as that of Woolliams and Meuwissen (21), could be more optimal.

The coefficients of variation considered were the same as those considered by Nicholas (15) and cover a substantial range of degrees of risk aversion. In particular, when national breeding schemes are considered, risk aversion seemed reasonable. But, in a competitive situation, a breeding firm may lag behind its competitors, and an increase of the lag may not harm the firm much (e.g., the firm has to buy improved breeding stock anyway). In this case, breeding schemes with high risk may be preferred, because there is at least a chance to outperform the competition, and a criterion regarding the probability of getting a response larger than a certain level is appropriate, as was studied by Dekkers (4). Woolliams and Meuwissen (21) considered risk preference by putting a reward on prediction error variance.

The assertion of Nicholas (15) that the variance of the response forms a more stringent restriction than the rate of inbreeding, may be criticized for two reasons. First, he used the coefficient of variation of the difference be-

tween a selection and a control line for this purpose. The variation of the genetic levels of both lines contributed equally to this variance, but, in practice, no control lines exist, and only the variation of the selected line remains. Thus, variances and effective sizes are halved. Second, Nicholas (15) made implicit use of the formula $V(\Delta G) = 2\Delta F\sigma_a^2$, which assumes random mating and overpredicts the variance of the response when selection is carried out. Meuwissen (13) showed that, in some cases, variance of response was only 42% of that predicted from this formula. Hence, effective sizes required for coefficients of variation of .32 and .16 may be as little as .21 (= .5 × .42) of those given by Nicholas (15), resulting in sizes of 21 and 84, respectively. This population size is of the same order of magnitude as the population size required for annual rates of inbreeding of .005, namely, effectively 27 animals per generation (generation interval is 3.67 yr in an adult MOET scheme). However, this effective size of 27 animals per generation is still rather small to prevent deterioration of fitness (14). In conclusion, if the coefficient of variation constraint is .32, the annual rate of inbreeding is probably about .005, but lower rates of inbreeding might be required. If the coefficient of variation constraint is .16, rates of inbreeding will be substantially lower and, probably, acceptable.

If more than 8 offspring could be obtained from an elite cow, genetic gain increased by 2% (Tables 2 and 4), which is substantially less than the increases predicted in other studies (11, 16, 21) because the variance of the response was restricted here. The variance of the response probably depended more on the actual number of animals selected than on the proportion of the animals selected and thus the selection differentials. Hence, larger nuclei could have both a sufficiently large number of elite cows selected and a small proportion selected, which yield high selection differentials. Therefore, the benefit of high reproductive rates of elite cows increases as nucleus size increases.

In general, the results showed that breeding schemes changed considerably when a more stringent coefficient of variation was required, but genetic gains were not much reduced. To test the robustness of the optimized schemes, the open nucleus schemes with 8 offspring per elite cow (see Tables 2 and 3) were altered such that all the elite cows selected from the base (DN(B)) were rejected (e.g., did not pass quarantine requirements) and all nucleus progeny were obtained from the selected nucleus animals by increased MOET efforts. This modification hardly affected rates of genetic gain, but coefficients of variation of those animals were increased from .160 to .290 and from .315 to .397 for the schemes in which the coefficient of variation constraint = .16 and .32 schemes, respectively. Hence, the coefficients of variations are much more sensitive to deviations from the optimal schemes than the rates of gain. Because coefficients of variation are sensitive to variations in the breeding scheme and genetic gains are only slightly reduced by decreasing the required coefficients of variation, recommendations to adopt breeding schemes with low coefficients of variation, e.g., .16, seem reasonable.

Strong restrictions on variance of responses and implicitly on rates of inbreeding favored progeny testing of young bulls and open nuclei, which contrasted with the MOET nucleus schemes of Nicholas and Smith (16) that were obtained when higher variances of responses were allowed. High reproductive rates of elite cows favored closed nuclei, which led to hybrid MOET schemes (2). Generally, progeny testing of young bulls proved to be the most effective method to reduce variances of responses and maintaining high rates of genetic gain.

ACKNOWLEDGMENTS

The authors are grateful for the funding from the Ministry of Agriculture Fisheries and Food, the Milk Marketing Board of England and Wales, and the Meat and Livestock Commission.

REFERENCES

1 Bulmer, M. G. 1971. The effect of selection on genetic variability. Am. Nat. 105:201.
2 Colleau, J. J. 1985. Genetic improvement by ET within selection nuclei in dairy cattle. Genet. Sel. Evol. 17:499.
3 De Boer, I.J.M., and J.A.M. van Arendonk. 1992. Optimizing the additive response to selection adjusted for effects of inbreeding in a closed dairy cattle nucleus assuming a large number of gametes per

female. A.F.G. Green and J.H.J. van der Werf, ed. Page 26 *in* Proc. Workshop Dep. Anim. Breeding, Wageningen Agric. Univ. Rep. No. 28. Wageningen, The Netherlands.

4 Dekkers, J.C.M., and G. E. Shook. 1990. Economic evaluation of alternative breeding programs for commercial artificial insemination firms. J. Dairy Sci. 73: 1902.

5 Greve, T., and V. Madison. 1991. In vitro fertilization in cattle: a review. Reprod. Nutr. Dev. 31:147.

6 Henderson, C. R. 1982. Best linear unbiased prediction in populations that have undergone selection. Page 191 *in* Proc. World Congr. Sheep and Beef Cattle, Massey Univ. Vol. 1. R. A. Barton and W. C. Smith, ed. Dunmore Press Ltd., Palmerston North, New Zealand.

7 Henderson, C. R. 1984. Application of Linear Models in Animal Breeding. Univ. Guelph, Guelph, Canada.

8 Hill, W. G. 1974. Prediction and evaluation of response to selection with overlapping generations. Anim. Prod. 18:117.

9 Hill, W. G. 1977. Variation in response to selection. Page 511 *in* Proc. Int. Conf. Quant. Genet. E. Pollak, O. Kempthorne, and T. B. Bailey, ed. Iowa State Univ. Press, Ames.

10 Johnson, D. L. 1977. Variance-covariance structure of group means with overlapping generations. Page 511 *in* Proc. Int. Conf. Quant. Genet. E. Pollak, O. Kempthorne, and T. B. Bailey, ed. Iowa State Univ. Press, Ames.

11 Meuwissen, T.H.E. 1991. The use of increased female reproductive rates in dairy cattle breeding schemes. Anim. Prod. 52:21.

12 Meuwissen, T.H.E. 1991. Reduction of selection differentials in finite populations with a nested full-half sib family structure. Biometrics 47:195.

13 Meuwissen, T.H.E. 1991. Expectation and variance of genetic gain in open and closed nucleus and progeny testing schemes. Anim. Prod. 53:133.

14 Meuwissen, T.H.E., and J. A. Woolliams. 1993. Effective sizes of livestock populations to prevent decline in fitness. A. F. Groen and J.H.J. van der Werf, ed. Page 26 *in* Proc. Workshop Dep. Anim. Breeding, Wageningen Agric. Univ. Rep. No. 28, Wageningen, The Netherlands.

15 Nicholas, F. W. 1989. Incorporation of new reproductive technology in genetic improvement programs. Evolution and Animal Breeding. W. G. Hill and T.F.C. Mackay, ed. CAB Int., Wallingford, England.

16 Nicholas, F. W., and C. Smith. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. Anim. Prod. 36:341.

17 Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1989. Numerical Recipes. Cambridge Univ. Press, Cambridge, UK.

18 Quinton, M., C. Smith, and M. E. Goddard. 1992. Comparison of selection methods at the same level of inbreeding. J. Anim. Sci. 70:1060.

19 Rendel, J. M., and A. Robertson. 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. J. Genet. 50:1.

20 Woolliams, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. Anim. Prod. 49:1.

21 Woolliams, J. A., and T.H.E. Meuwissen. 1993. Decision rules and variances of response in breeding schemes. Anim. Prod. 56:179.

22 Woolliams, J. A., and C. Smith. 1988. The value of indicator traits in the genetic improvement of dairy cattle. Anim. Prod. 46:333.

23 Woolliams, J. A., and I. Wilmut. 1989. Embryo manipulation in cattle breeding and production. Anim. Prod. 48:3.

24 Wray, N. R., and R. Thompson. 1990. Prediction of rates of inbreeding in selected populations. Genet. Res. 55:41.

Paper 19

RESPONSE VERSUS RISK IN BREEDING SCHEMES

T.H.E. Meuwissen
ID-DLO Institute for Animal Science and Health
PO Box 501, 3700 AM Zeist, Netherlands

J.A. Woolliams
Roslin Institute 'Edinburgh'
Roslin, Midlothian EH25 9PS, Scotland

SUMMARY

Increased rates of gains in modern breeding schemes are often at the expense of the risk of a breeding scheme. There is a strong relationship between inbreeding and variance of response, such that addressing one of these two components of risk usually suffices. A simple relationship between $V(\Delta G)$ and ΔF is given and can be used to determine whether the rate of inbreeding or the variance of response should be restricted. For some typical examples, obtaining acceptable variances of responses formed a more stringent restriction on the effective population size than the rate of inbreeding.

Selection methods, that reduce variance of response or rates of inbreeding are discussed. Most methods put a cost on variance of response or rate of inbreeding and the effect of varying the cost factor is studied. Others, restrict variance of response or rates of inbreeding directly. There are many similarities between the methods and the results indicate almost invariably that large reductions of variance of response and rates of inbreeding can be combined with small reductions in genetic gain.

INTRODUCTION

Recent increases in genetic gains of breeding schemes are mostly achieved at the expense of the risk of the breeding plans. For instance, BLUP (Best Linear Unbiased Prediction; Henderson, 1984) uses family information to increase the accuracy of estimated breeding values. Use of family information increases the probability of selection of relatives, because they come from the same (good) family, and thus increases inbreeding and variance of response. Similarly, in MOET (Multiple Ovulation and Embryo Transfer) schemes genetic gains are mainly increased by reduction of generation intervals (Nicholas and Smith, 1983), but the latter results in selection of fewer animals per generation, more generations per unit of time, and selection of young untested animals based on family information. All these elements increase inbreeding and variance of response, which are two components of the risk of a breeding scheme.

Inbreeding leads to random drift of gene frequencies of the population and, consequently, loss of good genes, which reduces the genetic variance and thus future selection responses. If gene effects are dominant, inbreeding depression may occur for traits under selection and for other traits. Inbreeding depression is generally regarded as being largest for traits associated with fitness such as fertility and disease resistance (Falconer, 1989).

Variance of selection response leads to loss due to uncertainty in planning and investment in breeding schemes (Woolliams and Meuwissen, 1993). A higher response than expected leads

to extra benefits, but these may be smaller than the losses from restoring and correcting gains which occur if gains are lower than expected.

When genetic gains of breeding plans are maximized, attention should be paid to the risk of the breeding schemes. Breeding schemes with high rates of gain and low risks may be preferred over schemes with very high gains and high risks. Several methods to account for the risk of breeding schemes, while maximizing genetic gains, have been reported in the literature and will be reviewed here. But first we will consider the relationship between the two components of risk: inbreeding and variance of response.

VARIANCE OF RESPONSE AND INBREEDING

In the situation of a monoecious population without selection, the variance of the genetic mean of the randomly selected animals is σ_a^2/n_s, where σ_a^2 = the genetic variance and n_s = the number of animals selected. If the actual population size is much larger than n_s, the variance of the mean of the next generation is also σ_a^2/n_s. Hence, the response of random selection is on average (of course) zero and has a variance of σ_a^2/n_s. Further, σ_a^2/n_s equals the drift variance $2\Delta F\sigma_a^2$, where ΔF = the rate of inbreeding, which is $1/2n_s$ in a monoecious randomly selected population.

Selection affects the variance of the selection response through (Hill, 1977): i) the selected animals have a smaller genetic variance than the population genetic variance; ii) the population genetic variance will be reduced by selection (Bulmer, 1971); and iii) the increased variance of the genetic variance due to selection. Effect i) may be accommodated by considering the animals that exceed the selection threshold and have a genetic variance of σ_a^2 $(1-r^2k)$, where r = the accuracy of selection, and k = the relative phenotypic variance reduction due to selection. Under truncation selection k = i(i-x), where i = intensity of selection and x = the standardized truncation point. Effect ii) can be simply accommodated by accounting for the effect of selection on σ_a^2 and r using Bulmer's (1971) formulas. After some initial generations of selection, the population genetic variance will stabilize at $\sigma_{a\infty}^2$. Hence, the variance of the selection response becomes:

$$V(\Delta\bar{G}) = 2\Delta F\sigma_{a\infty}^2 (1-r_\infty^2 k), \qquad [1]$$

where ΔG = rate of genetic gain.

A computer simulation of a breeding scheme was conducted to test Formula [1]. In each generation, 10 male and 50 female animals were selected from 100 male and 100 female candidates. Phenotypic (mass) selection or BLUP selection was applied for 10 discrete generations. The initial heritability and phenotypic variance of the trait were 0.25 and 1, respectively. Genetic variance was reduced by linkage disequilibrium, but not by inbreeding. Over the last 5 generations and with mass selection: $\sigma_{a\infty}^2 = .208$, $r_\infty^2 = .217$, $\Delta F = .0169$, $V(\Delta G) = .0060$ from the simulations and $V(\Delta G) = .0059$ from Formula [1]. The formula for drift with random selection yields $2\Delta F\sigma_{a0}^2 = .00845$, which is an overestimate of 41%. With BLUP selection: $\sigma_{a\infty}^2 = .202$, $r_\infty^2 = .300$, $\Delta F = .0326$, $V(\Delta G) = .008$ from the simulations and $V(\Delta G) = .010$ from Formula [1]. Hence, Formula [1] is accurate for mass selection but overpredicts $V(\Delta G)$ with BLUP selection. Another derivation of $V(\Delta G)$, which is provided by Woolliams and Meuwissen (1994), sorts out the long term effects of selection on ΔF and $V(\Delta G)$ and yields more accurate predictions.

CONSTRAINTS ON VARIANCE OF RESPONSE AND/OR INBREEDING

Nicholas (1989) suggested the following constraints for an animal breeding plan: i) rates of inbreeding at around 0.5% per yr should be acceptable; and ii) coefficients of variation of the selection response should not be larger than 5-10% over a 10 yr period of selection. With larger coefficients of variation, realized responses will differ frequently and markedly from the expected response, which is not acceptable in practical breeding plans. Nicholas' constraints are also used here. A more detailed analysis of the effective sizes that are required in livestock populations is provided by Meuwissen and Woolliams (1994a).

Table 1 shows effective population sizes that are required to achieve rates of inbreeding of 0.5%/yr and coefficients of variation of 10%. With such constraints on the variance and inbreeding, the effective size that was required to meet the constraint on the coefficient of variation was mostly larger than that due to the inbreeding constraint. This would in particular be the case, if a coefficient of variation of 5% was required, which quadruples the required effective sizes. Hence, for these typical examples, optimization of breeding schemes with a constraint on the variance of response will mostly lead to schemes with acceptable rates of inbreeding. This was also concluded by Nicholas (1989), when considering adult MOET schemes.

If the optimization of a breeding scheme results in a scheme with very short generation intervals (about 1 yr) or selection is very accurate and intensive in all paths, the constraint on the rate of inbreeding is more stringent. Hence, in these situations, optimization with a constraint on the rate of inbreeding will lead to schemes with acceptable rates of inbreeding and coefficients of variation.

Table 1. The effective population sizes per generation required for a coefficient of variation of the selection response of 10% over 10 yr of selection, $N_e(CV_{10})$, and a rate of inbreeding of 0.5% per yr $N_e(F)$. If a coefficient of variation of 5% is required, effective population sizes are $4*N_e(CV_{10})$.

		$N_e(CV_{10})$								
		p=.05			p=.2			p=.5		
L	$N_e(F)$	r=.4	r=.6	r=.8	r=.4	r=.6	r=.8	r=.4	r=.6	r=.8
1	100	13	5	2	28	10	4	88	34	15
3	33	38	14	5	84	31	12	265	101	44
6	17	76	27	10	168	61	24	529	202	87

(L = generation interval in yr; r = accuracy of selection; p = fraction selected)

TRADING OFF RESPONSE VERSUS ITS VARIANCE

In the literature, several papers tried to find breeding schemes with acceptable variances of response and high rates of gain. Standard methods for dealing with uncertainty are developed in utility theory and for Bayesian decision rules. In the following these methods are applied to trade off genetic gain versus its variance.

Utility theory

Meuwissen (1991) used utility theory to trade off mean and variance of response. According to utility theory, the expected utility of breeding schemes should be maximized. First, a utility

function should be specified, which gives the utility of some realized selection response. Meuwissen assumed a quadratic utility function, mainly because higher order terms would require higher moments than the variance of the response, which are generally unknown. Hence, the utility function was $U(\Delta G) = \Delta G - a\,\Delta G^2$, where a = constant specifying the risk aversion. The breeding scheme with the highest expected utility, $E(U(\Delta G)) = E(\Delta G) - a\,(V(\Delta G) + E^2(\Delta G))$, was preferred.

The value of a may be obtained from an analysis of the profit of a breeding scheme as a function of its genetic gain, when utility equals profit. But, such an economical analysis is complicated and utility may express more than profit alone. Since utility does not decrease with ΔG, the utility function is only valid for $\Delta G < 1/2a$. This provides a maximum value for a: $a_{max} = 1/(2\max(\Delta G))$, where $\max(\Delta G)$ is the highest response that can be realized by the alternative breeding schemes. Meuwissen (1991) assumed that the probability that ΔG exceeded $E(\Delta G) + 2\sigma(\Delta G)$ was negligible. A drawback of this approach is that the utility function depends on the breeding schemes considered, while it is should depend only on risk aversion considerations.

Conventional and modern progeny testing schemes, and open and closed nucleus schemes were compared by Meuwissen (1991). The ranking of the breeding schemes based on $E(\Delta G)$ was almost the same as that on $E(U(\Delta G))$. However, the closed scheme had 3% more response than the open nucleus scheme, but had a lower utility because of its 46% higher $\sigma(\Delta G)$. Hence, if schemes did not differ much in $E(\Delta G)$ but substantially in $V(\Delta G)$, the scheme with the lowest $V(\Delta G)$ had the highest utility.

Bayesian selection rules

Woolliams and Meuwissen (1993) considered Bayes selection rules to select groups of animals within a round of selection and given their estimated breeding values, the vector $\hat{\mathbf{u}}$, and the prediction error variance matrix **PEV**. Let **c** be the vector of contributions of the animals to the selected group, i.e. $c_i = 1/n$ for selected animals and $c_i = 0$ otherwise. A profit function was defined as:

$$P(\mathbf{c}) = \mathbf{c}'\hat{\mathbf{u}} + f(|\mathbf{c}'\mathbf{u} - \mathbf{c}'\hat{\mathbf{u}}|) + g(\mathbf{c}'\mathbf{u} - \mathbf{c}'\hat{\mathbf{u}}) + h(\mathbf{c}'\hat{\mathbf{u}} - \mathbf{c}'\mathbf{u}),$$

where **u** = vector of true breeding values, f(x) = loss due to uncertainty of the mean of the true breeding values of the selected group, g(x) = extra profit when **u** was underestimated by $\hat{\mathbf{u}}$, h(x) = extra loss when **u** was overestimated by $\hat{\mathbf{u}}$. Further, $f(x) = g(x) = h(x) = 0$ for $x \leq 0$. The functions f, g, and h may be obtained from an economical analysis. Woolliams and Meuwissen assumed linear and quadratic forms for f, g and h.

Bayesian decision rules maximize the expected profit under the posterior distribution of the breeding values, which has mean $\hat{\mathbf{u}}$ and variance **PEV**. With linear forms, $f(x) = \alpha x$, $g(x) = \beta x$ and $h(x) = \gamma x$, the expected profit is:

$$E(P(\mathbf{c})) = \mathbf{c}'\hat{\mathbf{u}} + .8\,(\alpha + \tfrac{1}{2}\beta + \tfrac{1}{2}\gamma)\sqrt{(\mathbf{c}'\mathbf{PEV}\mathbf{c})} = \mathbf{c}'\hat{\mathbf{u}} + a\sqrt{(\mathbf{c}'\mathbf{PEV}\mathbf{c})},$$

where .8 is the expectation of the absolute value of a random deviate from the distribution N(0,1). When covariances between breeding value estimates are neglected, selection is for $\hat{u}_i + a\sqrt{PEV_{ii}}$, which implies selection for a percentile of the posterior distribution of the breeding value. For instance, if a = -1 and distributions are normal, it can be seen from standard normal distribution tables that selection is for the 16th percentile.

For quadratic functions $f(x) = \alpha x^2$, $g(x) = \beta x^2$ and $h(x) = \gamma x^2$, the expected profit is:

$$E(P(c)) = \mathbf{c'\hat{u}} + (\alpha + \tfrac{1}{2}\beta + \tfrac{1}{2}\gamma)\ \mathbf{c'PEVc} = \mathbf{c'\hat{u}} + b\ \mathbf{c'PEVc}, \qquad [2]$$

where b is usually negative implying a cost on prediction error variance. A quadratic profit function favours animals with low prediction error variances relatively more than a linear profit function. However, given the group of animals selected under the quadratic profit function (more precisely: when their realized $\mathbf{c'PEVc} = k$ is known), a linear profit function can be derived which selects the same animals (Woolliams and Meuwissen, 1993). Also, the same group of animals would have been selected if the estimated breeding value of the selected group was maximized under the constraint $\mathbf{c'PEVc} = k$.

In an adult MOET scheme, Woolliams and Meuwissen (1993) found that the probability of co-selection of full sibs was reduced with increasing values of b. For one particular value of b, the prediction error variance was about equal to that of an adult MOET scheme with the constraint that only one animal per sib-ship is selected (see Nicholas and Smith, 1983). But, this quadratic profit function yielded 12% more genetic gain than the breeding scheme with the constraint selection of one animal per sib-ship.

When applied to open nucleus dairy cattle breeding schemes, linear and quadratic profit functions could reduce coefficients of variation of genetic gain by about 50% while the reduction in genetic gain was only 4 -8%. The reduction of the variance of response was mainly achieved by increasing the generation intervals of sires from 2.5 to about 4.5 yr. When the generation interval was 4.5 yr, selection of a substantial proportion of the sires was based on progeny test results.

In practice linear profit functions may be simply adopted by selecting for, e.g, the 25%-ile of the distribution of the breeding values, which is $\hat{u}_i + x_{0.25}\sqrt{PEV_{ii}}$, where $x_{0.25}$ is negative and defined by $P(x \leq x_{0.25} | x \sim N(0,1)) = 0.25$. This approach would neglect prediction error covariances. Woolliams and Meuwissen (1993) suggested to publish next to expected breeding values, i.e., the 50%-ile, the 25%-ile and, for breeders that like risk, the 75%-ile. Selection for several generations for the 75%-ile was shown to yield on average marginally more response than selection for the estimated breeding values: the reduction in genetic variance due to selection was smaller, which leads to higher responses in the next generation.

Constraints on the coefficient of variation of the genetic gain

When the genetic gain of a breeding scheme is maximized, some (implicit) constraint on the variance of response has to be adopted, otherwise only one animal per path would be selected leading to an unviable breeding scheme. Usually, the number of animals selected was fixed, but variances of response and rates of inbreeding can vary substantially even with a fixed number of animals selected. Therefore, Meuwissen and Woolliams (1994b) constrained the variance of response directly, while optimizing the number of animals selected per path and the generation intervals. This resulted in a large combinatorial optimization problem, because the number of animals selected per age class and path had to be optimized. The problem was solved by developing fast deterministic procedures to calculate rates of gain and their variances and by using the simulated annealing optimization algorithm (see Press et al, 1989).

This technique was used to optimize open nucleus dairy cattle breeding schemes while constraining the coefficient of variation to 5 and 10% over a 10 year period of selection (see Nicholas, 1989). If the number of offspring per donor cow was 8 and the coefficient of

variation was restricted to 10%, the resulting genetic gain was about 5% higher than that obtained by Meuwissen (1991), who optimized a similar open nucleus schemes with fixed numbers of animals selected. When the coefficient of variation was restricted to 5%, i.e., it was halved, the genetic gain reduced only by 2.3%. This low coefficient of variation was obtained by increasing the number of sires selected from 6 to 19 and selecting sires mainly from the older age classes. Selection of older animals reduces the variance of response because i) there are fewer selection rounds per unit of time; ii) selection is usually more accurate; and iii) the number of animals selected per generation increases. In the female selection path, the number of donor cows selected from commercial herds increased, i.e., the nucleus became more open.

If very large numbers of offspring per donor cow can be obtained by new reproductive techniques, the number of offspring per donor cow should be optimized in a breeding scheme. This optimization led to breeding schemes with only 1% more genetic gain than those with 8 offspring per donor cow, if the coefficient of variation was 10%. This figure probably increases with nucleus size, which was 256 male and 256 female animals born per yr. Further, the optimized open nucleus scheme was completely closed. If the coefficient of variation was restricted to 5%, genetic gain reduced by 4% and four progeny tested bulls were selected per yr. The latter breeding scheme was similar to the hybrid MOET schemes of Colleau (1985).

In all previous breeding schemes, each sire was mated to one or more dams. If a dam is mated to one or more sires, breeding schemes with a coefficient of variation of 10% are very similar to those with a sire mated to several dams, except that the number of sires and dams selected are exchanged. However, if the coefficient of variation was reduced to 5%, the genetic gains decreased by 10% and there were 30 donor cows selected with a generation interval of 3.3 yr. In the situation with selection of fewer males than females, the coefficient of variation could be reduced by selecting relatively few old progeny tested bulls, i.e., their genetic lag is compensated by the high accuracy of selection. With fewer females than males selected, old females were not very accurately tested and thus many young females were selected. The latter resulted in the larger reduction in response.

TRADING OFF RESPONSE VERSUS INBREEDING

Reduced rates of inbreeding at equal selection differentials

With truncation selection, the numbers of offspring of selected animals are intended to be equal. Toro and Nieto (1984) suggested to select more animals than with truncation selection and to have a varying number of offspring per animal such that the selection differential is the same as that with truncation selection. Thus, the best animal obtains the largest number of offspring and the worst animal the smallest. The number of offspring per animal was chosen such that the effective population size, $N_e = 1/\Sigma f_i^2$, was maximal, where f_i = the fraction of the offspring produced by parent i. The optimal values of f_i were obtained from quadratic programming, which maximized N_e under the restriction $\Sigma f_i \hat{u}_i = s$, where $\hat{u}_i$ = estimated breeding value of animal i and s = selection differential with truncation selection. In a small example, the effective population size could be increased from 5 to almost 6 by this approach. Unfortunately, this study did not consider the effects of more generations of selection on inbreeding. The probability of selection of offspring of poor parents may be very low because i) their estimated breeding value is low and ii) they have few relatives, which reduces their accuracy of selection. Hence, it is unlikely that poor parents have grand-offspring although they

were selected as parents and long term effective population sizes might be considerably smaller than $1/\Sigma f_i^2$.

A cost on the coefficient of kinship

Brisbane and Gibson (1993) suggested selection of groups of animals based on their average breeding value estimates minus a cost for their average genetic relationship, i.e., $\mathbf{c'\hat{u}} + a\,\mathbf{c'Ac}$ was maximized for $\mathbf{c}$, where a = a (negative) cost factor and $\mathbf{A}$ = the relationship matrix of the selection candidates. The similarity of this criterion with formula [2] becomes clear if we write $\mathbf{PEV} = (\mathbf{Z'R^{-1}Z} + \mathbf{A^{-1}}\sigma_a^{-2})^{-1}$, where $\mathbf{Z}$ = design matrix relating records to animals, and $\mathbf{R}$ = variance matrix of environmental effects of records. Now, if environmental variances are small, formula [2] reduces to $\mathbf{c'\hat{u}} + b\sigma_a^2\,\mathbf{c'Ac}$, which is equal to the criterion of Brisbane and Gibson if their cost factor $a = b\sigma_a^2$. Hence, Formula [2] select groups of animals with much information on the breeding values and low relationships, whereas $\mathbf{c'\hat{u}} + a\,\mathbf{c'Ac}$ considers only the relationships of the selected group.

At similar rates of inbreeding, selection for $\mathbf{c'\hat{u}} + a\,\mathbf{c'Ac}$ yielded higher responses than i) restricting the number of full sibs selected; ii) omitting varying amounts of records of relatives from the breeding value estimation; iii) selection on biased estimated breeding values calculated by reducing the weights on relatives records. Also, by varying the value of a, large reductions in rates of inbreeding could be achieved while reductions in genetic gain were small. The value of a could not be obtained from a cost/benefit analysis like the Bayesian profit functions of Woolliams and Meuwissen (1993).

Restrictions on rates of inbreeding

A drawback of putting a cost on the coefficient of kinship is that the number of animals selected has to be predefined, while this has a large effect on the rate of inbreeding. Quinton and Smith (1994) used simulated annealing to select a group of animals with a predefined average relationship $\mathbf{c'Ac}$ and maximum $\mathbf{c'\hat{u}}$, but without predefining the numbers of animals selected. By increasing the required average relationship each generation to $2\Delta F$, the rate of inbreeding should be restricted to ΔF. It is still under investigation whether this predefined rate of inbreeding is actually achieved and how large the reductions in genetic gain are.

DISCUSSION

Some recently proposed methods to reduce risk in breeding schemes were discussed. More methods were proposed in the literature, but the general results are almost invariably the same: large reductions in risk are possible without decreasing genetic gain much. For instance, 50% reduction in $V(\Delta G)$ is accompanied by 4-8% reduction in $E(\Delta G)$ (Woolliams and Meuwissen, 1993).

Several classifications of the methods are possible:

i) Methods that address ΔF versus $V(\Delta G)$. Formula [1] shows a strong relationship between these components of risk. Using formula [1], the desired N_e based on $V(\Delta G)$ constraints can be compared to that based on inbreeding constraints. If the former is the largest, the variance constraint is more stringent and should be applied and *vice versa*. Table 1 showed for some typical examples, that constraining $V(\Delta G)$ led mostly to acceptable rates of inbreeding (Table 1). However, if generation intervals are short and/or selection is intense and accurate, constraining rates of inbreeding is more appropriate.

ii) Some methods maximize genetic gain while constraining $V(\Delta G)$ or ΔF, whereas others put a cost on $V(\Delta G)$ or ΔF (or related parameters). Using Lagrangian multiplier methods, it can be shown that both categories select the same animals, if the realized $V(\Delta G)$ (or ΔF) of the costs methods equals the $V(\Delta G)$ (or ΔF) that was required in the constrained maximization methods (Woolliams and Meuwissen, 1993).

An advantage of the constraint maximization methods is that the numbers of animals selected, which affect $V(\Delta G)$ or ΔF substantially, are optimized simultaneously with the other parameters. A further advantage is that $V(\Delta G)$ or ΔF are predetermined in these methods. When putting a cost on $V(\Delta G)$ or ΔF, the resulting $V(\Delta G)$ or ΔF is not predetermined. However, simulations of breeding schemes while varying the cost factor will reveal the desired cost factor.

iii) Strategies that apply to entire breeding schemes versus rules to select a group of animals given their breeding values, prediction error (co)variances and genetic relationships of the selection candidates. The latter methods may profit from chance fluctuations in breeding values, etc., and will therefore probably yield higher response rates at the same $V(\Delta G)$ or ΔF. Also, they are more flexible to adapt to a particular breeding scheme.

When compared to selection for BLUP estimated breeding values ($\hat{u}$), the presented methods will reduce the weight of the between family component of the $\hat{u}$ and increase that of the within family component. This will reduce the probability of co-selection of sibs and thus reduce inbreeding and variance of response. No formal optimization of this re-weighting has been attempted, except by Woolliams (1994) who proposes an optimal weighting of the within family deviations of all the ancestors of an animal into a selection criterion.

REFERENCES

BRISBANE, J.R. and GIBSON, J.P. (1993) J. Dairy Sci.(Abs.) 76, Suppl. 1, 292.
BULMER, M.G. (1971) Am. Nat. 105: 201-211.
COLLEAU, J.J. (1985) Genet. Sel. Evol. 17: 499-538.
FALCONER, D.S. (1989) Longman Group Ltd, Essex, UK.
HENDERSON, C.R. (1984) Canad. Catal. Publ. Data, Univ. of Guelph, Canada.
HILL, W.G. (1977) In Proc. Intern. Conf. Quant. Genet. eds. Pollak, E., Kempthorne, O. and Bailey, T.B., 511-519.
MEUWISSEN, T.H.E. (1991) Anim. Prod. 53: 133-141.
MEUWISSEN, T.H.E. and WOOLLIAMS, J.A. (1994a) Theor. Appl. Genet. (accepted)
MEUWISSEN, T.H.E. and WOOLLIAMS, J.A. (1994b) J. Dairy Sci. (accepted)
NICHOLAS, F.W. (1989) In Evol. Anim. Breed. eds. Hill, W.G. and Mackay, T.F.C., CAB Int., Wallingford, UK, 201-209.
NICHOLAS, F.W. and SMITH, C. (1983) Anim. Prod. 36: 341-353.
PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A. and VETTERLING, W.T. (1989) Cambr. Univ. Press, Cambridge, UK.
QUINTON, M. and SMITH, C. (1994) Paper in preparation.
TORO, M.A. and NIETO, B.M. (1984) Genet. Res. Camb., 44: 347-349.
WOOLLIAMS, J.A. (1994) Proc. 5th Wrld. Congr. Genet. Appl. Livest. Prod.
WOOLLIAMS, J.A. and MEUWISSEN, T.H.E. (1993) Anim. Prod. 56: 179-186.
WOOLLIAMS, J.A. and MEUWISSEN, T.H.E. (1994) Paper in preparation.

Paper 20

Genet Sel Evol (1994) 26, 517-535

Original article

Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle

B Villanueva [1], JA Woolliams [2], G Simm [1]

[1] *Scottish Agricultural College, West Mains Road, Edinburgh, EH9 3JG;*
[2] *Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, UK*

(Received 24 December 1993; accepted 18 May 1994)

Summary – A closed MOET (multiple ovulation and embryo transfer) nucleus scheme, with overlapping generations, was modelled for beef cattle by stochastic simulation. Selection was carried out for 25 years on a trait measurable in both sexes and with a heritability of 0.35. Different strategies to control the rate of inbreeding were investigated: 1) decreasing female selection intensity whilst keeping the number of donors constant; 2) culling selected animals after having been used for a period of time; 3) using more donors; 4) using factorial mating designs; and 5) selecting on modified indexes. Comparisons among different schemes were made on the basis of equal number of transfers per year. Strategies 1, 2, and 3 reduced inbreeding but also reduced response. When the schemes were compared at the same level of inbreeding, culling of animals gave higher rates of genetic progress than decreasing selection intensity. Factorial designs decreased the rate of inbreeding by up to 19% in comparison with nested designs, with no effect on response. The most successful strategies were those that reduced the emphasis on family information in the selection criterion and especially selection on estimated breeding values obtained by BLUP (best linear unbiased prediction) using a deliberately increased heritability. With this method, it was possible to reduce inbreeding by up to 30% without affecting genetic progress. The reduction in inbreeding with different raised heritabilities averaged 42% and ranged from 26 to 61%. Under all the strategies studied to control inbreeding, proportional reductions in rates of inbreeding were always higher than those in genetic response.

beef cattle / breeding scheme / MOET / genetic gain / inbreeding

Résumé – Stratégies pour contrôler la consanguinité dans des schémas de sélection fermés avec transfert d'embryons chez les bovins à viande. *Un schéma de sélection fermé de bovins à viande, utilisant le système MOET (ovulation multiple et transfert d'embryon), et avec des générations imbriquées, a été soumis à un modèle de simulation stochastique. La sélection pendant 25 ans a porté sur un caractère mesurable dans les 2 sexes et d'héritabilité 0,35. Différentes stratégies pour contrôler le taux de consanguinité ont été examinées : i) réduction de l'intensité de sélection en sélectionnant un nombre plus*

grand de femelles, tout en maintenant un nombre constant de donneuses ; ii) élimination des animaux (donneuses ou pères) après une seule période d'évaluation (6 mois) ; iii) utilisation de plus de donneuses ; iv) utilisation de plans factoriels de croisement ; v) sélection selon des indices modifiés. Des comparaisons ont été faites entre les différents schémas, à nombre égal de transferts par an. Les stratégies iii), ii), i) conduisent à une réduction du taux de consanguinité, mais la réponse aussi est réduite. Quand on compare les différents schémas à niveau égal de consanguinité, l'élimination précoce des animaux donne un taux de progrès génétique plus élevé que la réduction de l'intensité de sélection. Les plans factoriels réduisent le taux de consanguinité d'une quantité pouvant aller jusqu'à 19% par rapport aux plans hiérarchiques, sans aucun effet sur les réponses. La stratégie qui donne les meilleurs résultats est la sélection sur les valeurs génétiques additives obtenues au moyen du BLUP en utilisant une héritabilité délibérément augmentée. Avec cette dernière méthode, la consanguinité est réduite jusqu'à 30% tandis que le progrès génétique reste constant. Une autre stratégie qui réduit le taux de consanguinité consiste à sélectionner sur un indice modifié pour diminuer la contribution de l'information familiale. Dans chacune des stratégies examinées pour contrôler la consanguinité, la réduction proportionnelle de la consanguinité a toujours été plus grande que celle de la réponse.

schéma de sélection / bovin à viande / ovulation multiple et transfert d'embryon / gain génétique / consanguinité

INTRODUCTION

Improved reproductive rates of females through multiple ovulation and embryo transfer (MOET) can lead to an increase in genetic response, due to increased selection intensities and reduced generation intervals. In the absence of the effects of inbreeding, Land and Hill (1975) indicated that the rates of genetic progress for growth rate in beef cattle could be doubled by using MOET in comparison with conventional schemes. Gearheart *et al* (1989) extended these results to different selection criteria and heritabilities and also found increases in genetic responses from MOET. These studies predicted response after a single generation of selection. Stochastic simulations, which have accounted for factors which influence medium or long-term responses, have shown that these theoretical predictions substantially overestimated the advantage of MOET schemes (Wray and Simm, 1990).

Comparisons among alternative breeding schemes have usually been made on the basis of expected rates of genetic progress. However, in practice, breeding schemes are operated with restrictions on rates of inbreeding, either implicitly or explicitly, to limit its negative effects (loss of genetic variation and inbreeding depression). One of the main drawbacks of MOET nucleus schemes is the increased rates of inbreeding resulting from their small population size. Faster inbreeding occurs with any selection scheme involving between-family selection (Robertson, 1961). The larger family sizes created by MOET amplifies this effect. Wray and Simm (1990) have shown that when comparing MOET with conventional beef breeding schemes at the same level of inbreeding, the advantage of MOET in genetic response was reduced to around 50%.

Several strategies have been proposed to control the rate of inbreeding in selection programmes (*eg*, Toro and Perez-Enciso, 1990). All of these strategies have either

direct or indirect effects on restricting the magnitude of the variance of family size and the expected relationship of long-term genetic contribution of ancestors with their breeding value (Wray and Thompson, 1990). For a given number of transfers, the variance of family size is least when all females contribute equally to descendants in subsequent generations. Increasing the opportunity of a female to be used as a donor decreases the variance of family size. This can be achieved by increasing the number of donors used in a period and by culling donors immediately following a designated number of flushes.

Best linear unbiased prediction (BLUP) is generally accepted as the optimum procedure for genetic evaluation. By using all information on relatives, the accuracy of estimating the breeding value is increased. However, selection methods in which the accuracy of prediction is gained by using ancestral information, can lead to higher rates of inbreeding due to the higher probability of selecting related animals (Robertson, 1961). Dempfle (1975) showed that, in the long term, selection within families could give higher selection response than individual selection, mostly due to the maintenance of genetic variability resulting from the increase in effective population size. He showed that, with selection on phenotypes, the advantage of within-family selection increases when the heritability is high and with large families. MOET schemes, with the use of BLUP, benefit progress, in the short term, by increasing family sizes and accuracies. By using a selection criterion in which the weight given to family information is reduced, inbreeding rates might be decreased without greatly affecting response.

Once the selection decisions have been made, the choice of the mating system can also affect the rates of genetic progress and inbreeding. Factorial mating designs, in which each dam is mated to more than one sire, were proposed by Woolliams (1989) for MOET breeding schemes to reduce rates of inbreeding with no loss in response.

In this paper, different strategies to control inbreeding are investigated through Monte-Carlo simulation of a closed MOET beef nucleus herd.

METHODS

Description of simulations

Basic scheme

A MOET nucleus scheme with overlapping generations was simulated for beef cattle. An additive infinitesimal genetic model was assumed. True breeding values of unrelated base animals (9 males and 18 females) were obtained from a normal distribution with mean zero and variance (σ^2_A) 0.35. Phenotypic values were obtained by adding a normally distributed environmental component with mean zero and variance 0.65. Thus, initial heritability was 0.35. Equal numbers of animals of 2, 3 and 4 years of age were simulated. To mimic selection for beef trait, it was assumed that the trait under selection was recorded in both sexes at around 400 d of age (between 385 and 415 d), at the end of a performance test. Selection was carried out for 25 years. The number of breeding males and females (donors) was constant over years and equal to the number of base males and females (9 males

and 18 females). Animals were genetically evaluated twice every year (evaluation period = 6 months). An estimate of breeding value (EBV) was obtained for each animal using an individual animal model-BLUP. The only fixed effect included in the model was the overall mean. All the information available at the time of evaluation was used to obtain the EBVs. Males and females with the highest EBVs were selected. There were no restrictions on the number of sires or dams selected from any one sibship. In the absence of the culling policies described below, animals were selected irrespective of whether they had been selected in previous periods. Animals not selected were culled from the herd.

Values for reproductive parameters (minimum age of donors, frequency of collection and proportion of calves per transfer) were taken from Luo *et al* (1994) and represent the current realistic situation in embryo technologies. Each donor was flushed 3 times in each evaluation period (embryo collections were carried out every 2 months). The number of transferable embryos collected was obtained from a negative binomial distribution (Woolliams *et al*, 1994). The mean number of transferable embryos per flush and per donor was 5.1, with a coefficient of variation of 1.25 and repeatability of 0.23. These values were obtained from analyses of extensive data on embryo recovery (Woolliams *et al*, 1994). Thus, the average number of embryo transfers per year was around 550. All calves were born from embryo transfer, *ie* there were no calves from natural matings. Embryos transferred survived until birth with probability 0.55 and the sex ratio was expected to be 1:1 (sex was assigned at random with probability 0.5). Males were assumed capable of breeding at 12 months of age and females at 15 months of age. At all ages after birth, individuals were subject to a mortality rate that varied with age. The maximum age of the animals was 15 years. Selected donors and sires were randomly mated according to a nested mating design (each donor was mated to the same sire in consecutive flushes, within an evaluation period). Each sire was used the same number of times.

After year zero, true breeding values of the offspring born every year, were generated as

$$TBV_i = (1/2)(TBV_s + TBV_d) + m_i$$

where TBV_i, TBV_s and TBV_d are the true breeding values of the individual i, its sire and its dam, respectively, and m_i is the Mendelian sampling term. The Mendelian term was obtained from a normal distribution with mean zero and variance $(1/2)[1-(F_s+F_d)/2]\sigma^2_A$, where F_s and F_d are the inbreeding coefficients of the sire and dam, respectively. The inbreeding coefficients of the animals were obtained from the relationship matrix, using the algorithm proposed by Quaas (1976).

Alternative schemes

In order to control rates of inbreeding, several modifications of the basic scheme described in the previous section were considered. The different strategies studied are described below. Unless otherwise stated, the simulations were run as described for the basic scheme. Some combinations of different alternatives were also studied.

Selection intensity in females

The number of selected females in one period was increased from 18 (basic scheme) to 27, 36, 54, 72, 90, 108 and 144. In all cases, only 18 females, chosen at random from these selected females, were used as donors. In this way, the number of transfers was kept constant.

Limited use of selected parents

In a given period, each of the 18 donors was flushed 3 times and was then ineligible for further selection. Culling of males after use in one period was also examined.

Number of donors

At each evaluation period, 27 cows were selected and flushed twice. Thus, on average, the number of embryos was equal to that obtained with 18 donors flushed 3 times.

Mating design

A factorial mating design, in which donors were mated to different sires in consecutive flushes, was also considered. Each selected bull was used the same number of times and randomly assigned to donors.

Selection criteria

Three alternative selection criteria were studied. Firstly for each animal, a modified index ($IND1$) was computed as

$$IND1_i = EBV_i - \lambda_s EBV_s - \lambda_d EBV_d$$

where subscripts i, s and d refer to the individual, its sire and its dam and the EBV_s are those obtained from BLUP. Different values of λ_s and λ_d were used to explore the effects of a range of weights given to family information. Note that when $\lambda_s = \lambda_d = 1/2$, selection is based on the estimated Mendelian sampling component and so a form of within-family selection is practised. Animals with the highest index values were selected.

Secondly a selection criterion ($IND2$), which has been recently used by Grundy and Hill (1993), was evaluated. Individuals were selected according to their EBV obtained from BLUP using an artificially raised heritability (h^2_{AR}). Different values for h^2_{AR} were examined (from 0.5 to 0.9).

Finally, for each animal, a modified index ($IND3$) was computed as

$$IND3_i = EBV_i - |EBV_i \gamma F_i|$$

where subscript i refers to the individual; the EBV is that obtained from BLUP and F is the inbreeding coefficient. Different values for the factor γ were investigated. Again, selected animals were those with the highest index values. This index can

be seen as a method to achieve retrospective minimum coancestry matings. By penalizing individuals with high inbreeding coefficients in the selection decisions, matings of highly related animals are penalized retrospectively.

Comparison among breeding schemes

The basic scheme was used as a point of reference for comparisons. Average true breeding values (G_i) and inbreeding coefficients (F_i) of individuals born at the ith year were obtained. Rates of response between years j and i were calculated as $\Delta G_{i-j} = G_j - G_i$, where $j > i$. Rates of inbreeding were obtained every year as $\Delta F = (F_i - F_{i-1})/(1 - F_{i-1})$. Other parameters calculated in the simulations were: 1) genetic variance of animals born every year; 2) accuracy of selection (correlation between the true breeding values and selection criteria of the candidates for selection); 3) genetic selection differentials (difference between the mean values of selection criteria of selected individuals and candidates for selection) and selection intensities for males and females; 4) generation intervals (average age of parents when offspring are born) for males and females; and 5) variance of family sizes for male and female parents. To calculate the variance of family size, the cohort of calves born at year 11 was chosen (each year should be similar to any other after genetic parameters approach equilibrium). Let M_{11} and F_{11} represent, respectively, males and females born at year 11, which are selected to produce offspring at any time. The variance of family size for males was calculated as $\mathrm{Var}(n_m) + \mathrm{Var}(n_f) + 2\,\mathrm{Cov}(n_m, n_f)$, where n_m and n_f are, respectively, the number of male and female offspring of M_{11} that are selected at any time. The variance of family size for females was calculated in a similar way by counting offspring of F_{11} that are selected in successive years. Appropriate variances and covariances of family sizes were calculated at the end of each replicate. The number of replicates ranged from 20 to 50. Values presented are the average over all replicates.

The number of transfers per year was expected to be the same for all the schemes studied. The criteria for comparing different schemes were the rates of response and inbreeding at different time periods. The cumulative response and inbreeding at year 15 were also compared.

RESULTS

Selection intensity

Genetic responses and inbreeding coefficients obtained per year, for different female selection intensities, are shown in figure 1. The number of selected females initially varied from 18 to 144, although in all cases, only 18 females were used as donors. Rates of response decreased substantially after year 5 due to the decrease in genetic variance by linkage disequilibrium (Bulmer, 1971). This decrease in variance is greatest during the first generation of selection (selection of animals born from base animals starts at the third year) and then slowly approaches an equilibrium. After that, the change in genetic variance is due to inbreeding. Rates of inbreeding become approximately constant after year 15 (around 5 generations of selection). The same pattern of response and inbreeding over years was observed for all the

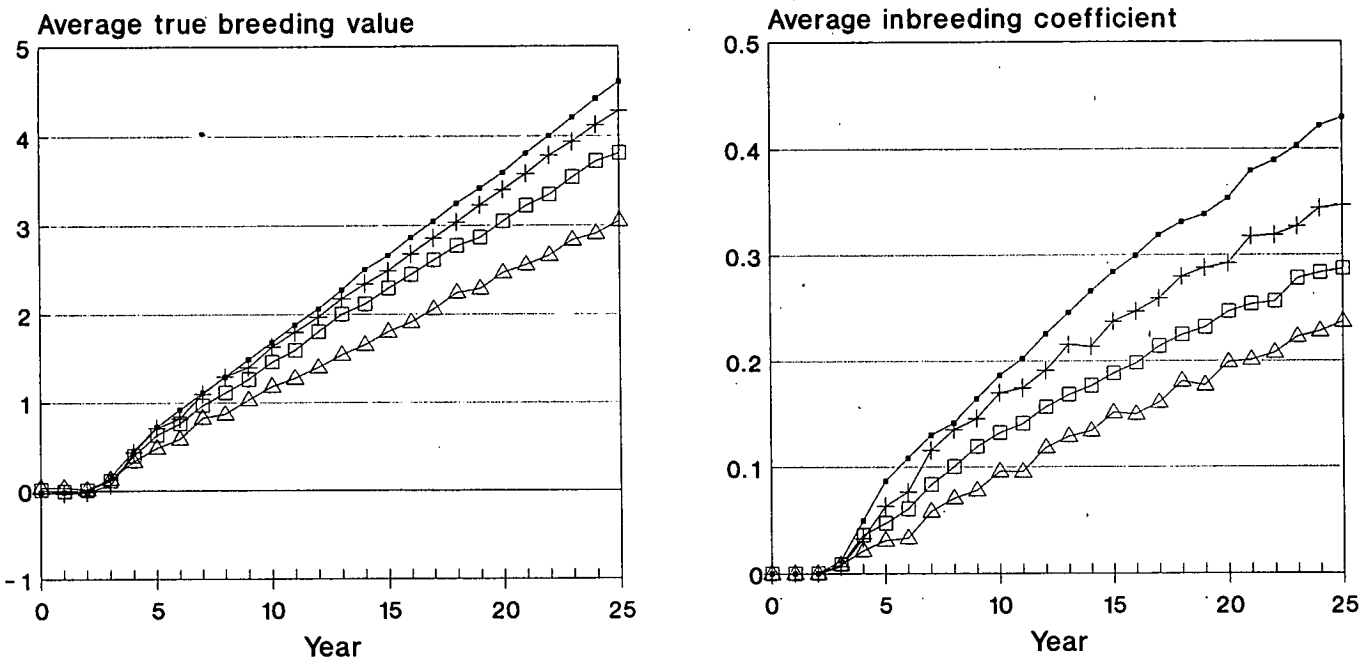


Fig 1. Change in genetic mean (phenotypic standard deviation units) and inbreeding coefficient (%) over years with different female selection intensities. The number of donors was 18 in all cases. ■: 18 selected; +: 27 selected; □: 54 selected; ▵: 144 selected.

schemes studied. For these reasons, 2 times periods were considered. Average rates of response from year 5 to 15 (ΔG_{5-15} and ΔF_{5-15}) and from year 15 to 25 (ΔG_{15-25} and ΔF_{15-25}) under different female selection intensities are shown in table I. Cumulative selection responses to year 25 (G_{25}) and average inbreeding coefficients at this year (F_{25}) are also presented. As expected, decreasing intensity of selection led to a decrease in rates of response and inbreeding. Decreasing selection intensity reduced rates of inbreeding (ΔF_{15-25}) by 34 to 58%, whereas rates of response (ΔG_{15-25}) were reduced by 7 to 36%, compared with the case where 18 females were selected. Rates of response and inbreeding were, in general, slightly higher in the early years (5−15) than in later years (15−25). Table II shows selection intensities and generation intervals obtained in the last time period for males (i and L) and females (i and L). Decreasing selection pressure in females led to a decrease in L (fewer donors are repeatedly used over successive periods). However, this was accompanied by a small increase in male generation interval, probably due to the slower genetic progress achieved. The average generation interval ranged from 2.94 to 3.09 years.

Limited use of selected parents, number of donors and mating design

Table III shows the effect of different culling policies, number of donors and mating designs on rates of response and inbreeding. Culling of females after each evaluation period reduced inbreeding but also reduced response. For the different number of donors and mating designs considered, the culling of females reduced the rate of inbreeding by 24–37%. Corresponding proportional reductions in response were lower (4–12%). When males were also culled from the herd after each period, there was, in general, a further reduction in inbreding rates. However, the rates of response were similar to those obtained when only females were culled. Although culling of males led to decreased generation intervals, there was no further reduction in the intensity of selection. Culling of animals resulted in a better strategy for decreasing inbreeding than reducing selection intensity (see also table I). That is, for the same level of inbreeding, there was a smaller reduction in genetic progress by culling animals than by reducing intensity of selection. Generation intervals for males and females obtained for the different schemes are shown in table IV. The values presented are averages from year 15 to 24. Culling of animals decreased generation intervals by around 16%.

Increasing the number of donors used from 18 (3 flushes per period) to 27 (2 flushes per period) led to reductions in inbreeding and in response. Differences in rates of inbreeding between schemes using 18 and 27 donors were smaller under the factorial mating design. For the different culling policies and mating designs considered, increasing the number of donors decreased rates of inbreeding by 2–38% and rates of response by 2–13%. Generation intervals were slightly increased by increasing the number of donors used (table IV).

The factorial mating design gave, in general, a slightly higher response (not statistically significant, $P < 0.05$) than the nested design and significantly lower rates of inbreeding (table III). When the number of donors was 18 and animals were allowed to be repeatedly selected (*ie* no culling), the factorial design reduced the rate of inbreeding by 19%. The average variance of family sizes after selection for female parents (over replicates) was 6.71 and 4.48 with nested and factorial designs,

Table I. Rates of genetic progress (phenotypic standard deviation units) and inbreeding (%) and their standard errors (se) for different female selection intensities under a nested mating design.

No of females selected	*Genetic progress*						*Inbreeding*					
	$\Delta G_{5\text{-}15}$	*(se)*	$\Delta G_{15\text{-}25}$	*(se)*	G_{25}	*(se)*	$\Delta F_{5\text{-}15}$	*(se)*	$\Delta F_{15\text{-}25}$	*(se)*	F_{25}	*(se)*
18**	0.194	(0.003)	0.194	(0.003)	4.60	(0.06)	2.33	(0.13)	2.15	(0.18)	42.93	(1.16)
27*	0.178	(0.005)	0.180	(0.005)	4.28	(0.07)	1.96	(0.25)	1.42	(0.17)	34.64	(1.80)
36*	0.179	(0.005)	0.170	(0.005)	4.15	(0.08)	1.68	(0.14)	1.28	(0.13)	31.69	(1.71)
54*	0.167	(0.004)	0.152	(0.004)	3.81	(0.07)	1.53	(0.10)	1.24	(0.12)	28.68	(1.18)
72*	0.146	(0.005)	0.148	(0.005)	3.43	(0.08)	1.32	(0.12)	1.32	(0.11)	25.93	(1.35)
90*	0.139	(0.004)	0.125	(0.004)	3.15	(0.08)	1.29	(0.10)	0.90	(0.11)	22.45	(1.30)
108*	0.139	(0.007)	0.132	(0.006)	3.16	(0.09)	1.37	(0.07)	1.04	(0.14)	23.96	(1.14)
144*	0.132	(0.004)	0.125	(0.003)	3.05	(0.08)	1.29	(0.09)	1.04	(0.13)	23.67	(1.20)

* 20 replicates; ** 50 replicates; the number of donors was 18 in all cases.

Table II. Selection intensities (i and i) and generation intervals (L and i , in years) for males and females and their standard errors (se) for different female selection intensities under a nested mating design.

No of females selected	i	*(se)*	i	*(se)*	L	*(se)*	L	*(se)*
18	1.36	(0.006)	1.63	(0.009)	3.18	(0.014)	3.00	(0.013)
27	1.15	(0.009)	1.65	(0.011)	2.93	(0.008)	3.04	(0.019)
36	0.98	(0.012)	1.65	(0.012)	2.82	(0.007)	3.07	(0.026)
54	0.63	(0.013)	1.67	(0.013)	2.72	(0.005)	3.17	(0.026)
72	0.34	(0.014)	1.67	(0.015)	2.68	(0.006)	3.20	(0.022)
90	0.10	(0.011)	1.66	(0.012)	2.67	(0.005)	3.24	(0.026)
108	0.01	(0.007)	1.65	(0.016)	2.66	(0.007)	3.32	(0.021)
144	−0.01	(0.005)	1.68	(0.015)	2.68	(0.006)	3.33	(0.036)

The number of donors was 18 in all cases.

respectively. Corresponding averages for the variance for male parents were 39.24 and 41.68, but there was enormous variation among replicates in these values. The variance of family size for males varied from 0 to 256 in the nested and from 0.5 to 174 in the factorial design. The efficiency of factorial designs for controlling inbreeding rates was smaller when 27 females were used as donors. There were no differences in generation intervals between mating designs (table IV).

Selection criteria

The rates of response and inbreeding, obtained by using different selection criteria, are presented in table V. Three different modified indexes (*IND*1, *IND*2 and *IND*3), as described above, were studied as alternatives to selection on BLUP breeding values. Males and females were culled after each selection period. For all schemes considered, the generation intervals ranged from 2.42 to 2.56 years for males and from 2.54 to 2.65 years for females. Selection on the index *IND*1 (table V) indicated that, by decreasing the contribution of family information, inbreeding levels were greatly reduced. The reduction in response was mostly due to a decrease in the accuracy of selection. Average accuracy from year 14 to 24 was 0.57 with BLUP and 0.46 with *IND*1 and $\lambda_s = \lambda_d = 1/2$. As would be expected, the decline in genetic variance was smaller with selection on the index. Average values from year 14 to 24 for the genetic variance ranged from 0.24 (BLUP) to 0.28 ($\lambda_s = \lambda_d = 1/2$). With culling, generation intervals were kept approximately constant (2.55 years for males and 2.65 years for females) by varying λ_s and λ_d. For values of $\lambda_s = \lambda_d = \lambda$, response decreased up to around 19% whereas inbreeding decreased up to 31% ($\lambda = 1/2$). For values of λ between 0.2 and 0.33, inbreeding decreased substantially whereas the change in response was very small. For higher values of λ, the decreases in inbreeding and response were notable. Figure 2 shows trends in rates of response and inbreeding obtained for different values of λ. It can be observed that rates of inbreeding are much more sensitive to the change in λ than rates of response.

Results obtained when λ_s and λ_d differ are also presented in table V. Genetic response (and inbreeding) was slightly higher when $\lambda_s > \lambda_d$ (when the weight given,

Table III. Rates of genetic progress (phenotypic standard deviation units) and inbreeding (%) and their standard errors (se) for different culling policies, number of donors and mating designs.

Mating design	*No of donors*	*Culling*		*Genetic progress*				*Inbreeding*			
				$\Delta G_{15\text{-}25}$	*(se)*	G_{25}	*(se)*	$\Delta F_{15\text{-}25}$	*(se)*	F_{25}	*(se)*
Nested	18	No	No**	0.194	(0.003)	4.60	(0.06)	2.15	(0.18)	42.93	(1.16)
		Yes	No**	0.184	(0.003)	4.49	(0.05)	1.58	(0.13)	39.00	(1.32)
		Yes	Yes**	0.185	(0.003)	4.37	(0.05)	1.22	(0.13)	35.13	(0.96)
	27	No	No*	0.186	(0.004)	4.42	(0.08)	2.06	(0.23)	38.10	(1.80)
		Yes	No*	0.164	(0.004)	3.99	(0.05)	1.27	(0.15)	30.01	(1.44)
		Yes	Yes*	0.174	(0.005)	3.85	(0.08)	0.75	(0.23)	30.02	(1.63)
Factorial	18	No	No**	0.188	(0.003)	4.66	(0.04)	1.73	(0.12)	36.31	(1.09)
		Yes	No*	0.181	(0.005)	4.29	(0.07)	1.31	(0.13)	31.12	(1.33)
		Yes	Yes*	0.186	(0.005)	4.41	(0.06)	1.15	(0.17)	30.11	(1.27)
	27	No	No*	0.184	(0.005)	4.43	(0.08)	1.69	(0.19)	33.35	(2.14)
		Yes	No*	0.173	(0.006)	4.05	(0.07)	1.13	(0.12)	27.20	(1.16)
		Yes	Yes*	0.161	(0.004)	3.78	(0.06)	0.79	(0.14)	25.66	(1.65)

* 20 replicates; ** 50 replicates.

Table IV. Generation intervals (years) for males and females and their standard errors (se) for different culling policies, number of donors and mating designs.

Mating design	*No of donors*	*Culling*		L	*(se)*	L	*(se)*
Nested	18	No	No	3.18	(0.014)	3.00	(0.013)
		Yes	No	2.64	(0.003)	3.04	(0.013)
		Yes	Yes	2.64	(0.003)	2.54	(0.004)
	27	No	No	3.30	(0.024)	3.12	(0.025)
		Yes	No	2.66	(0.004)	3.13	(0.028)
		Yes	Yes	2.66	(0.003)	2.61	(0.005)
Factorial	18	No	No	3.18	(0.011)	3.00	(0.011)
		Yes	No	2.64	(0.004)	3.01	(0.023)
		Yes	Yes	2.64	(0.004)	2.54	(0.006)
	27	No	No	3.32	(0.026)	3.14	(0.027)
		Yes	No	2.65	(0.004)	3.10	(0.025)
		Yes	Yes	2.67	(0.004)	2.62	(0.005)

in the selection criterion, to family information from the male side is smaller than that given to information from the female side) although difference between $\lambda_s > \lambda_d$ and $\lambda_s > \lambda_d$ were unclear. Also, there were no clear differences in components of response (selection intensity, accuracy, genetic variance and generation interval).

Results obtained when some artificially raised values for the heritability (h^2_{AR}) were used in the BLUP evaluations (*IND*2) are also presented in table V. The true heritability was 0.35. For values of h^2_{AR} equal to or smaller than 0.7, response was kept practically constant whereas the rate of inbreeding decreased by 26–42%. For values of h^2_{AR} greater than 0.7, response decreased by 4–6% whereas the rate of inbreeding decreased by 48–61%. Trends in rates of response and inbreeding can also be observed in figure 2, which shows that *IND*2 is more efficient than *IND*1 in controlling inbreeding. Of all schemes considered these were the most effective for decreasing inbreeding without affecting response.

When the modified index *IND*3, which penalizes individuals with high inbreeding coefficients in selection decisions, was used, there was no decrease in the rate of inbreeding. However, the response was affected.

DISCUSSION

The control of rates of inbreeding has become important in the design of breeding programmes since several procedures, introduced in the first instance to produce extra gains (MOET, BLUP), can in fact have a dramatic impact on inbreeding. These procedures can result in proportionally higher increases in rates of inbreeding than in rates of response compared to conventional schemes and mass selection. All the strategies evaluated for decreasing rates of inbreeding in a closed nucleus MOET herd were efficient in the sense that rates of inbreeding were reduced proportionally more than rates of response. The best strategy (to reduce inbreeding with little

Table V. Rates of genetic progress (phenotypic standard deviation units) and inbreeding (%) and their standard errors (se) for different criteria under a nested mating design with limited use of parents.

Selection criterion	*Genetic progress*				*Inbreeding*			
	ΔG_{15-25}	*(se)*	G_{25}	*(se)*	ΔF_{15-25}	*(se)*	F_{25}	*(se)*
BLUP**	0.185	(0.003)	4.37	(0.05)	1.22	(0.13)	35.13	(0.96)
*IND*1								
$\lambda_s = \lambda_d = 1/9^*$	0.183	(0.006)	4.36	(0.08)	1.16	(0.28)	35.17	(2.11)
$\lambda_s = \lambda_d = 1/8^{**}$	0.183	(0.003)	4.35	(0.04)	0.80	(0.18)	33.52	(1.28)
$\lambda_s = \lambda_d = 1/7^{**}$	0.184	(0.003)	4.33	(0.05)	1.30	(0.13)	34.62	(1.17)
$\lambda_s = \lambda_d = 1/6^{**}$	0.185	(0.003)	4.35	(0.04)	0.98	(0.14)	32.38	(1.09)
$\lambda_s = \lambda_d = 1/5^*$	0.179	(0.006)	4.27	(0.08)	0.83	(0.18)	28.88	(1.83)
$\lambda_s = \lambda_d = 1/4^*$	0.175	(0.005)	4.25	(0.06)	0.79	(0.17)	29.41	(1.38)
$\lambda_s = \lambda_d = 1/3^*$	0.175	(0.005)	4.13	(0.04)	0.69	(0.18)	27.65	(1.12)
$\lambda_s = \lambda_d = 1/2^*$	0.151	(0.006)	3.53	(0.04)	0.74	(0.17)	24.15	(0.79)
$\lambda_s = 1/2, \lambda_d = 0^*$	0.184	(0.005)	4.20	(0.08)	0.62	(0.33)	29.35	(1.41)
$\lambda_s = 0, \lambda_d = 1/2^*$	0.177	(0.005)	4.15	(0.08)	0.65	(0.14)	27.23	(1.07)
$\lambda_s = 1/2, \lambda_d = 1/8^*$	0.175	(0.004)	4.07	(0.06)	1.02	(0.20)	27.85	(1.27)
$\lambda_s = 1/8, \lambda_d = 1/2^*$	0.166	(0.006)	3.97	(0.05)	0.71	(0.24)	27.63	(1.52)
*IND*2								
$h^2_{AR} = 0.5^{**}$	0.182	(0.003)	4.39	(0.04)	0.90	(0.11)	27.67	(1.04)
$h^2_{AR} = 0.6^*$	0.185	(0.006)	4.23	(0.06)	0.84	(0.14)	24.38	(0.95)
$h^2_{AR} = 0.7^*$	0.187	(0.005)	4.33	(0.07)	0.71	(0.11)	22.75	(1.19)
$h^2_{AR} = 0.8^*$	0.175	(0.005)	4.19	(0.08)	0.63	(0.11)	19.27	(1.04)
$h^2_{AR} = 0.9^*$	0.170	(0.004)	4.11	(0.07)	0.48	(0.08)	17.23	(1.27)
*IND*3								
$\gamma = 1^*$	0.171	(0.005)	4.27	(0.07)	1.27	(0.20)	34.94	(1.56)
$\gamma = 2^*$	0.159	(0.007)	4.01	(0.09)	1.01	(0.27)	34.00	(1.63)

* 20 replicates; ** 50 replicates; the number of donors was 18.

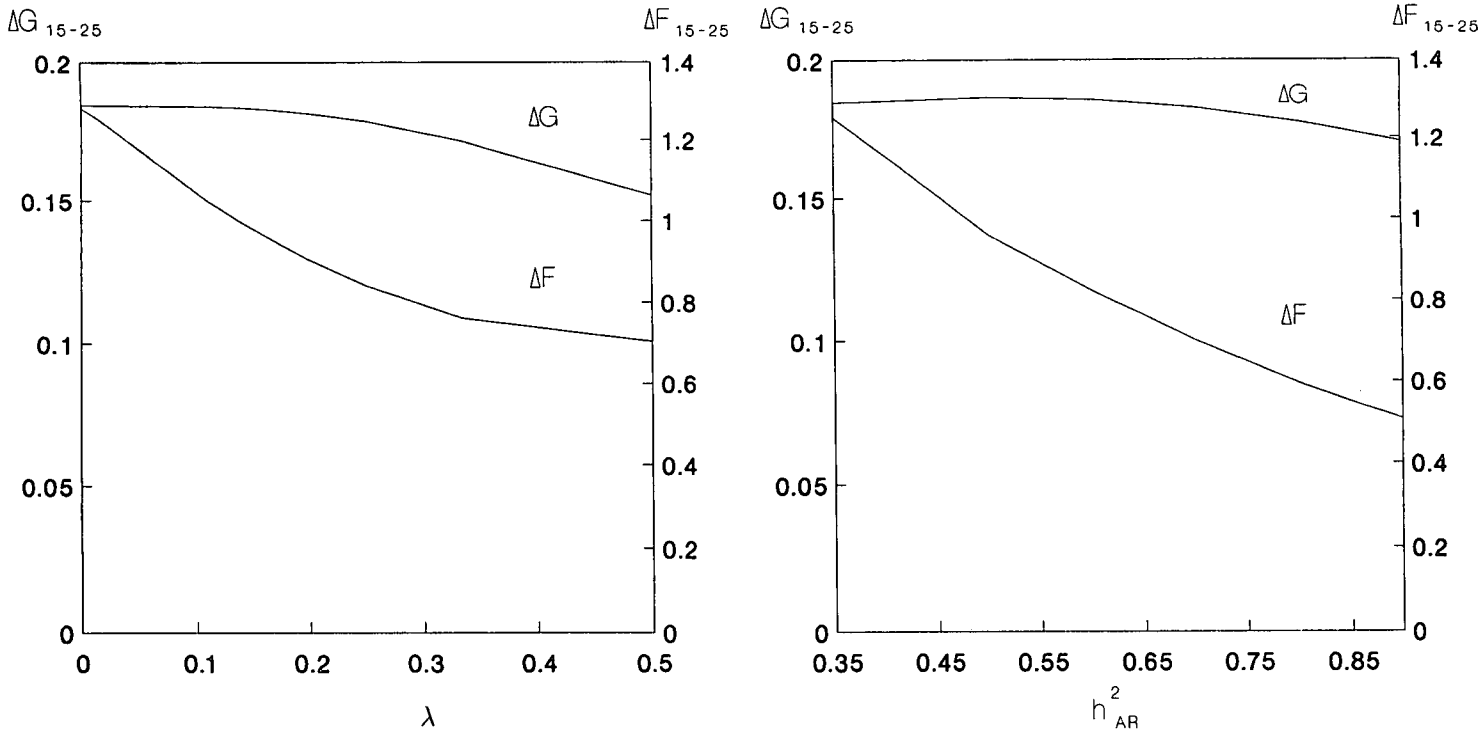


Fig 2. Rates of response (ΔG_{15-25}) and inbreeding (ΔF_{15-25}) obtained by using the modified indexes *IND*1 (with different values of $\lambda_s = \lambda_d = \lambda$) and *IND*2 (with different values of h^2_{AR}).

effect on response) was selection on modified indexes (especially *IND*2) in which the weight given to family information is reduced. Factorial designs were also capable of keeping gain constant and decreasing inbreeding although this decrease was smaller than with *IND*2. The other strategies led to clear reductions in response. Of these, the best was culling of animals after being used for a period of time since for a given level of inbreeding, the response was higher than that obtained using more donors or reducing selection intensity. Costs of implementing the different schemes for beef cattle would be similar if it were assumed that the most critical cost factor is the overall number of embryos transferred.

Rates of genetic progress and inbreeding greatly depend on parameters of embryo yield distributions. When previous studies on possible benefits from MOET in beef cattle (Land and Hill, 1975; Gearheart *et al*, 1989; Wray and Simm, 1990) were carried out, good estimates of the necessary reproductive parameters were not available. Now we have reliable estimates for these parameters. In the present simulations, parameters for embryo recovery and embryo transfer were obtained from an extensive literature review and a survey of experts (Luo *et al*, 1994) and analyses of data (Woolliams *et al*, 1994). The number of transferable embryos was obtained from a Poisson distribution whose parameter is distributed according to a gamma distribution. Thus, extra variation in embryo yield was introduced in comparison with a strict Poisson (with constant parameter). Without control this will influence the rates of inbreeding observed through additional variation in family size.

The rate of inbreeding increases, in general, with the variance of family size. An indirect method for decreasing this variance is to decrease the intensity of selection. Culling of animals from the herd after being flushed a given number of times and flushing more donors, can directly reduce the variance of family sizes. The results presented show that with these strategies, although inbreeding was reduced, response was also affected. If schemes are compared at the same level of inbreeding, limiting the use of selected parents (*ie* culling animals) could give better results than unrestricted selection. One of the advantages of BLUP is that selection can be made across generations. In the light of these results, this advantage could be arguable when a longer term response is considered.

Woolliams (1989) proposed the use of factorial mating designs in MOET schemes either to increase response while keeping rates of inbreeding unchanged, or to decrease rates of inbreeding while keeping response constant. Previous simulation studies (Ruane, 1991; Stranden *et al*, 1991; Toro *et al*, 1991) restricted the number of sons (or daughters) from a full-sibship eligible for selection. With more full-sibships produced with factorial designs, selection intensity (and consequently response) were increased with no additional inbreeding. In the situation considered here, where selection intensities are maintained, factorial designs are expected to give the same genetic progress as nested designs but with lower rates of inbreeding. These predictions were consistently found in all simulations where 18 donors were used and rates of inbreeding were decreased by up to 19%. Increasing the number of donors to 27 leads to a reduction in the variance of family size and the advantage of factorial designs is reduced. For the same reason, culling of animals under the factorial design led, in general, to smaller reductions in inbreeding than under the nested design (table III).

Selection on BLUP breeding values is expected to give the highest response in the short term. These higher responses are, however, accompanied by higher rates of inbreeding. In addition, the use of family information in genetic evaluation increases the correlation among estimated breeding values causing lower than expected selection differentials and response (Hill, 1976). Finally, the decline in genetic variance due to linkage disequilibrium (Bulmer, 1971) increases with the accuracy of selection. Therefore, responses obtained with methods that predict breeding values more accurately decrease proportionally more than responses from less accurate methods. Selection on modified indexes that decrease the family contribution seems to be a promising method to control inbreeding since genetic progress appears very robust to different weights given to family information. The use of *IND*1 can substantially decrease inbreeding with a small change in response. These results have also been found by Verrier *et al* (1993) using an index equivalent to *IND*1 when $\lambda_s = \lambda_d$. Grundy and Hill (1993) have utilized an alternative approach, initially proposed by Toro and Perez-Enciso (1990), to reduce the weights attached to family information. This involves artificially raised heritabilities in the BLUP evaluations in order to reduce the weight given to the family mean and therefore reduce co-selection of relatives. They showed that, in this way, inbreeding rates can be reduced with only a small loss in response. This also agrees with the results obtained by Toro and Silio (1993). The procedure has been evaluated in the present study (*IND*2) showing that methods exist which can reduce inbreeding by more than 40% with little effect on response.

Although both *IND*1 and *IND*2 are based on the same principles (reducing the weight given to family information), their efficiencies differ, with *IND*2 givin better results. *IND*2 not only reduces the weights given to pedigree information but also affects the evaluations of the sire and the dam. Genetic progress can be viewed as the covariance of long-term contributions and Mendelian sampling terms (Woolliams and Thompson, 1994). The breeding value of an individual is the sum of the Mendelian sampling term specific to that individual, plus the average breeding value of its parents. The breeding value of the parents can also be decomposed into Mendelian sampling components and the average of the breeding values of the respective parents. This decomposition can be carried out for each generation, back to the base generation. Thus, the estimated breeding value of an individual is a sum of estimated Mendelian sampling terms weighted by $1/2^t$ for an ancestor occurring t generations previously. To reduce inbreeding the expected short-term gain must be sacrificed to some degree (maybe small) and this may be seen as a decision on which information on genetic merit to ignore. The further back from the current generation, the greater the potential for inbreeding relative to the amount of information obtained. With *IND*2, by increasing the heritability, the weights attached to information from previous generations are progressively reduced each generation and therefore inbreeding is greatly reduced with little effect on response (the difference in weighting of information with respect to standard BLUP is greater for generations furthest from the current generation). With *IND*1, whilst extra weight is being given to the Mendelian sampling terms in the current generation, Mendelian sampling terms of all previous generations are weighted according to BLUP weights. Consequently, the reduction in inbreeding is less than that obtained

with *IND*2 and there is a higher reduction in response to obtain a given reduction in inbreeding.

The results obtained for *IND*2 assume a single value of the true heritability (0.35). Simulations were run for schemes with different levels of heritability and results showed, in all cases, the high efficiency of *IND*2. By using $h^2_{AR} = 0.5$ in a scheme with heritability 0.2, the rate of inbreeding was reduced by 41% whereas response was only reduced by 5%. Corresponding reductions in rate of inbreeding and response by using $h^2_{AR} = 0.8$ in a scheme with heritability 0.5 were 38 and 1%. The method was also efficient in a larger scheme with 36 donors (1 100 transfers/year) and 9 sires. By using $h^2_{AR} = 0.7$ (where the true heritability was 0.35) and limiting the use of parents, the rate of inbreeding was reduced by 55% whereas response was reduced by 4%.

Previous studies comparing different selection procedures have considered non-overlapping generations. Results presented in this paper for different selection criteria (table V) correspond to cases where animals are culled from the herd after being used during one period. This situation is therefore similar to that considered in previous studies in the sense that comparisons are only among contemporaries. It could be argued that the efficiency of *IND*2 for decreasing the rate of inbreeding at minimal cost in gain would be reduced when comparisons are made across animals born in different generations. Simulations were run with no culling of animals and the results show that the method is also very efficient when generations overlap. For example, when the heritability was artificially raised to 0.7, the rate of inbreeding decreased by 23% but the rate of response was not greatly affected (in comparison with the basic scheme in which the heritability used was 0.35). One disadvantage of this method is that breeding values, and therefore genetic trends, are wrongly estimated. However, unbiased trends and estimates of fixed effects could be obtained by running the evaluations with the unbiased estimate of heritability. Selection would be carried out on the data corrected for the fixed effects but reanalysed using the artificially raised heritability.

In this study, selection and mating procedures have been analysed separately. When the best selection and mating strategies (*IND*2 and factorial mating design) were combined in one single scheme, the rate of inbreeding was reduced by around 52% whereas response was not substantially affected. Thus, the change in rate of inbreeding from using both strategies simultaneously was similar to the sum of the changes from using both strategies separately.

With the exception of factorial mating designs and selection on *IND*2, in general, the strategies considered here will also decrease the response to selection, for the number of generations considered in this paper (around 8). If selection was to be carried out for long enough, within family selection could give higher responses than mass selection (Dempfle, 1975; Verrier *et al*, 1993). Moreover results for Quinton *et al* (1992) have indicated that, in the long term, phenotypic selection can result in higher genetic gains than BLUP selection. Verrier *et al* (1993) have shown that, after 30 generations of selection, selection on a modified index (*IND*1 with $\lambda = 0.25$), can give higher responses than BLUP selection if the size of the population is small. Some of the modified heritabilities would have an even more dramatic effect. However, in practice, selection for the same objective is rarely practised for this length of time in closed populations.

The results discussed here assume additive genetic models, which account for the loss of genetic variance due to inbreeding. Models including inbreeding depression need further investigation. Variation in response, which also depends on the effective population size, is also an important parameter to be considered in comparisons among breeding programmes (Nicholas, 1989). Theoretical work is needed to compare objectively the different procedures suggested to control rates of inbreeding and to find optimum schemes setting the genetic gains from selection against the losses due to inbreeding.

ACKNOWLEDGMENTS

We are grateful to Drs R Lewis, R Thompson, NR Wray and BJ McGuirk for useful comments and to the Ministry of Agriculture, Fisheries and Food, the Milk Marketing Board of England and Wales and the Meat and Livestock Commission for funding this project. SAC also receives financial support from the Scottish office Agriculture and Fisheries Department.

REFERENCES

Bulmer MG (1971) The effect of selection on genetic variability. *Am Nat* 105, 201-211

Dempfle L (1975) A note on increasing the limit of selection through selection within families. *Genet Res (Camb)* 24, 127-135

Gearheart WW, Smith C, Teepker G (1989) Multiple ovulation and embryo manipulation in the improvement of beef cattle: relative theoretical rates of genetic change. *J Anim Sci* 67, 2863-2871

Grundy B, Hill WG (1993) A method for reducing inbreeding with best linear unbiased prediction. *Anim Prod* 56, 427 (abstr)

Hill WG (1976) Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* 32, 889-902

Land RB, Hill WG (1975) The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Anim Prod* 21, 1-12

Luo ZW, Woolliams JA, Simm G (1994) An assessment of present and future effectiveness of embryological techniques in ruminants. *Ag Biotech News Infor* 6,13N-18N

Nicholas FW (1989) Incorporation of new reproductive technology in genetic improvement programmes. *In: Evolution and Animal Breeding* (WG Hill, FC MacKay, eds), CBA International, Wallingford, UK, 203-209

Quaas RL (1976) Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32, 949-953

Quinton M, Smith C, Goddard ME (1992) Comparison of selection methods at the same level of inbreeding. *J Anim Sci* 70, 1060-1067

Robertson A (1961) Inbreeding in artificial selection programmes. *Genet Res (Camb)* 2, 189-194

Ruane J (1991) The effect of alternative mating designs and selection strategies on adult multiple ovulation and embryo transfer (MOET) nucleus breeding schemes in dairy cattle. *Genet Sel Evol* 23, 47-65

Stranden I, Maki-Tanila A, Mantysaari EA (1991) Genetic progress and rate of inbreeding in a closed adult MOET nucleus under different mating strategies and heritabilities. *J Anim Breed Genet* 108, 401-411

Toro M, Perez-Enciso M (1990) Optimization of selection response under restricted inbreeding. *Genet Sel Evol* 22, 93-107

Toro M, Silio L (1993) Comparacion de estrategias de optimizacion de respuesta con restriccion de consanguinidad. *ITEA* 12, 295-297

Toro M, Silio L, Perez-Enciso M (1991) A note on the use of mate selection in closed MOET breeding schemes. *Anim Prod* 53, 403-406

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

Woolliams JA (1989) Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Anim Prod* 49, 1-14

Woolliams JA, Thompson (1994) A theory of genetic contributions. *In: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, vol 19, Guelph, Canada, 127-134

Woolliams JA, Luo ZW, Villanueva B *et al* (1994) Analysis of embryo yields and ovulation rates. *J Agric Sci (Camb)*, 123 (in press)

Wray NR, Simm G (1990) The use of embryo transfer to accelerate genetic improvement in beef cattle. *In: Proceedings of the 4th World Congress on Genetics Applied to Livestock Production, vol XV*, Edinburgh, UK, 315-318

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

Paper 21

Genet Sel Evol (1995) 27, 347-363

Original article

The effect of improved reproductive performance on genetic gain and inbreeding in MOET breeding schemes for beef cattle

B Villanueva [1], JA Woolliams [2], G Simm [1]

[1] *Scottish Agricultural College, West Mains Road, Edinburgh, EH9 3JG;*
[2] *Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, UK*

(Received 9 November 1994; accepted 6 March 1995)

Summary – The effect of improved reproductive techniques on genetic progress and inbreeding was investigated in MOET (multiple ovulation and embryo transfer) schemes for beef cattle. Stochastic simulation was used to model a closed scheme with overlapping generations. Selection was on a trait measured in both sexes, with heritability 0.35, and was carried out for 25 years. The number of breeding animals was 9 sires and 18 donors. Embryo production was modelled using a Poisson distribution with the parameter distributed according to a gamma distribution. The mean number of transferable embryos per flush and per donor was 5.0, with a coefficient of variation of 1.28 and repeatability between flushes of 0.22. This model was compared with models used in previous studies (fixed number of embryos per flush or variable number of embryos but with zero repeatability between flushes). The coefficient of variation and the repeatability of embryo yield influenced inbreeding rates. The rate of inbreeding was underestimated by up to 17% when variability of embryo production was ignored. Without a constraint on the number of calves born per year, improved success rates for embryo collection and embryo transfer technologies led to notable increases in genetic progress. However, the rate of inbreeding was also increased with improved techniques. When the number of calves born per year was fixed, genetic progress was maintained but inbreeding rates were substantially reduced (by up to 11%) with improved techniques due to the opportunity of equalizing family sizes. There was no benefit from sexed semen with constrained number of calves per year.

beef cattle / MOET / embryo / genetic gain / inbreeding

Résumé – Effet de l'amélioration des techniques de reproduction sur le progrès génétique et sur la consanguinité dans des schémas MOET pour bovins à viande. *Notre investigation avait pour but d'étudier l'effet de techniques de reproduction améliorées sur le progrès génétique et sur la consanguinité, dans le cadre de schémas MOET (ovulation multiple et transfert d'embryon) pour les bovins à viande. Grâce à une simulation stochastique, un schéma fermé a été modélisé avec générations chevauchantes. La sélection a été effectuée pendant une période de 25 ans, sur un caractère mesurable dans les 2 sexes, dont*

l'héritabilité était de 0,35. Les nombres de reproducteurs mâles et de donneuses étaient de 9 et 18 respectivement. La production d'embryons a été modélisée en utilisant une distribution de Poisson dont le paramètre avait une distribution gamma. Le nombre moyen d'embryons transférables recueillis par collecte et par donneuse était de 5,0 avec un coefficient de variation de 1,28 et avec une répétabilité de 0,22 entre collectes. Ce modèle a été comparé avec d'autres modèles utilisés dans des études antérieures (qui utilisaient un nombre déterminé d'embryons par collecte, ou un nombre variable d'embryons mais avec une répétabilité nulle entre collectes). Le coefficient de variation et la répétabilité de la production d'embryons influencent le taux de consanguinité. Si on ne tient pas compte de la variabilité de la production d'embryons, la sous-évaluation du taux de consanguinité peut atteindre 17%. Sans contrainte sur le nombre de naissances de veaux par an, un plus grand pourcentage de réussite dans la collecte d'embryons et l'amélioration des technologies de transfert contribuent ensemble à augmenter considérablement le progrès génétique. Cependant, l'amélioration des techniques a aussi pour effet d'augmenter le taux de consanguinité. Quand le nombre de veaux nés par an est fixé, le progrès génétique peut être maintenu, tout en réduisant le taux de consanguinité (jusqu'à 13%), en employant les techniques améliorées, à cause de la possibilité d'égaliser la taille des familles. Il n'y a aucun bénéfice à utiliser du sperme sexé quand le nombre de veaux par an est fixé.

bovin à viande / schéma à ovulation multiple et transfert d'embryon (MOET) / embryon / gain génétique / consanguinité

INTRODUCTION

The value of multiple ovulation and embryo transfer (MOET) in breeding schemes for increasing genetic gain has been widely studied in dairy cattle (see review by Dekkers, 1992; Ruane and Thompson, 1991) and to a lesser extent in beef cattle (Land and Hill, 1975; Gearheart *et al*, 1989; Keller *et al*, 1990; Wray and Simm, 1990) and sheep (Smith, 1986; Wray and Goddard, 1994). Early studies concentrated on extra genetic progress expected with MOET. More recent studies have also considered possible risks associated with the use of MOET techniques. By greatly increasing the numbers of progeny to be produced by individuals, genetic progress can be improved due to increased intensities of selection. However, the extra gains can be accompanied by increased inbreeding since fewer parents contribute to the next generation. The adverse effects of inbreeding (loss of genetic variability, loss of predictability of genetic gain and inbreeding depression) should be taken into account when optimum schemes for genetic improvement using reproductive technologies are investigated.

One of the main shortcomings in earlier studies was the assumption of constant family size, or the assumption of a variable family size, but with no correlation between the number of embryos produced in successive recoveries. In fact there is a wide range in the size of families following MOET and analyses of MOET data have indicated a non-zero repeatability of embryo production (*eg*, Lohuis *et al*, 1993; Woolliams *et al*, 1995). The increase in the variance of embryo yield can lead to increased rates of inbreeding and reductions in genetic gain. Recently, Woolliams *et al* (1995) have proposed a mathematical model to describe the distributions of embryo yields observed in practice. The model includes repeatability (*ie* the

assumption of zero correlation between flushes is removed) and describes very accurately the number of embryos obtained per donor and per flush. Villanueva *et al* (1994) have used this model in a simulation study to investigate different strategies for controlling rates of inbreeding in MOET breeding schemes for beef cattle. With current values of parameters describing success rates of reproductive technologies, rates of inbreeding were very high for schemes with 18 donors and 9 sires, even when the most efficient strategies for controlling inbreeding were used (factorial mating designs and selection on best linear unbiased prediction (BLUP) breeding values assuming an inflated heritability). In this paper we investigate rates of progress and inbreeding obtained when different models for simulating embryo production are utilized.

Advances in embryo manipulation techniques have been rapid in the past few years and these are likely to continue. One of the main problems in the practical application of embryo transfer in breeding programmes using superovulation is the high variability among donors in the number of embryos collected. This produces a high variance in family size which in turn leads to a high variance in the numbers selected from each family (and, therefore, high inbreeding). Research is being addressed at reducing this variability and increasing the mean number of embryos per collection. Embryo survival rates and frequency of collection are also likely to be improved. Luo *et al* (1994) have given both pessimistic and optimistic predictions for future success rates of embryological techniques. The effect that improved future success rates for embryo recovery and embryo transfer could have on rates of response and inbreeding is investigated in this paper. Also, the techniques for sexing of embryos or semen are already used on a small scale. Semen and embryo sexing may become commercial in the near future and so the value of sexing of semen to increase genetic progress is also examined. Hence, the results are expected to be useful in identifying those advances in reproductive technologies which are likely to be of most value in breeding schemes.

METHODS

Description of simulations

The stochastic model to simulate a MOET nucleus scheme for beef cattle has been described in detail by Villanueva *et al* (1994). The trait under selection was assumed to be recorded in both sexes and around 400 d of age (between 385 and 415 d), at the end of a performance test. The trait was simulated assuming and additive infinitesimal model with an initial heritability of 0.35. The nucleus was established with 9 males and 18 females of 2, 3 and 4 years of age. The number of animals in each age group was approximately the same. These unrelated individuals constituted the base population. True breeding values of base population animals were obtained from a normal distribution with mean zero and variance (σ^2_A) 0.35 (different age groups had the same genetic mean). Phenotypic values were obtained by adding a normally distributed environmental component with mean zero and variance 0.65.

Selection was carried out for 25 years. The number of breeding males and females was constant over years and equal to the number of base males and females (9 sires and 18 donors). Animals were genetically evaluated twice a year (evaluation

period = 6 months). Estimated breeding values (EBVs) were obtained using an individual animal model BLUP. The overall mean was the only fixed effect included in the model. Males and females with the highest EBVs were selected and randomly mated according to a nested design. Each sire was used the same number of times in 1 evaluation period. Animals were selected irrespective of whether they had been selected in previous periods and animals not selected were culled from the herd.

True breeding values of the offspring born every year, were generated as

$$TBV_i = (1/2)(TBV_s + TBV_d) + m_i$$

where TBV_i, TBV_s and TBV_d are the true breeding values of the individual i, its sire and its dam respectively, and m_i is the Mendelian sampling term. The Mendelian term was obtained from a normal distribution with mean zero and variance $(1/2)[1 - (F_s + F_d)/2]\sigma_A^2$, where F_s and F_d are the inbreeding coefficients of the sire and dam, respectively. The inbreeding coefficients of the animals were obtained from the additive relationship matrix.

Values for reproductive success rates (parameters of embryo yield, frequency of embryo collection and survival rate of transferred embryos) were varied in different schemes. The number of transferable embryos collected per flush and per cow was obtained from a Poisson distribution whose parameter was distributed according to a gamma distribution (Woolliams *et al*, 1995). This model is described in the next section (*Model 1*). Different values for the mean number of transferable embryos per flush and per donor, the coefficient of variation and repeatability of embryo yield, the frequency of flushing and the embryo survival rate were considered. Current values were obtained from analyses of extensive data on embryo recovery (Woolliams *et al*, 1995). Potential future values were obtained from a survey of international experts in reproductive technologies (Luo *et al*, 1994). All calves were born from embryo transfer, *ie* there were no calves from natural matings. The survival to birth of a transferred embryo was assigned at random with different probabilities in different schemes (0.55, 0.65 or 0.75). The sex of the embryos was also assigned at random with probability 1/2 of obtaining a male (expected sex ratio ♂/♀ = 1:1) for most schemes. In order to evaluate the possible benefit of using sexed semen, the ratio was changed to 1:2 and 1:3. In these cases, the probability of obtaining a male was 1/3 and 1/4, respectively. Males were assumed capable of breeding after being performance tested. The minimum age of donors was 15 months. At all ages after birth, individuals were subject to a mortality rate that varied with age. Survival probabilities from birth to 3 weeks, 6 months and 2, 5, 10 or 15 years were 0.98, 0.97, 0.96, 0.93, 0.86 and 0.00, respectively. Thus, the maximum age of the animals was 15 years. Survival rates were assumed to be the same for both sexes.

Models for embryo production

In the present study, the number of embryos produced per flush and per donor was generated using the model proposed by Woolliams *et al* (1995). In order to investigate the effect of including extra variation in embryo production on rates of response and inbreeding this model was compared with models used in previous studies (fixed number of embryos per collection or variable number of embryos per

collection but with zero repeatability between flushes). Four different models were analysed.

Model 1

The model proposed by Woolliams *et al* (1995) generates the number of embryos produced from a negative binomial distribution (Poisson distribution whose parameter is distributed according to a gamma distribution). The number of embryos collected from the ith donor in the jth flush was sampled from a Poisson distribution whose parameter λ_{ij} was sampled from a gamma distribution with shape parameter β_i and scale parameter v. In that way a correlation between the number of embryos produced in successive flushes of a cow is included in the model. The natural logarithm of β_i (parameter specific for each donor) was sampled from a normal distribution with mean μ and variance σ^2. The logarithm of β_i is taken to avoid negative numbers. The maximum value of λ_{ij} was set to 30. Let y_i be the number of transferable embryos obtained at the jth collection. Then the expected value and variance of y_i are $E(y_i) = \beta_i v$ and $\mathrm{Var}(y_i) = \beta_i v(1 + \beta_i^2)$, respectively (Woolliams *et al*, 1995). In order to explore the effect of changing these key parameters, a simulation program was written to simulate embryo production using this model. The number of donors simulated was 100 000 and the number of flushes was 3 for each donor. The repeatability was calculated as $R = \sigma_B^2/(\sigma_B^2 + \sigma_W^2)$, where σ_B^2 is the variance in embryo production among donors and σ_W^2 is the variance among flushes (within donors). The coefficient of variation was calculated as $CV = (\sigma_B^2 + \sigma_W^2)^{1/2}/MEAN$, where $MEAN$ is the overall mean of embryos collected per flush and per donor. The estimates of σ_B^2 and σ_W^2 were obtained from an analysis of variance of simulated data. Current values for embryo production (Luo *et al*, 1994) correspond to the following parameter values: $\mu = 1.46$, $\sigma^2 = 0.4$ and $v = 1.0$. These values led to a mean number of transferable embryos per flush and per donor of 5.0, with a coefficient of variation of 1.28 and repeatability of 0.22.

Model 2

The number of embryos collected was obtained in the same way as described in *Model 1* but now the logarithm of β_i was sampled from a normal distribution with mean μ and variance zero. Since parameter β_i is a constant, there is no variability among donors and the repeatability of embryo production is zero. The values for the parameters of the distributions were $\mu = 1.61$, $\sigma^2 = 0.0$ and $v = 1.0$. These values lead to the same mean number of embryos collected as in *Model 1* but to a lower coefficient of variation ($CV = 1.09$, $R = 0.00$).

Model 3

The number of embryos collected per flush and per donor was generated by sampling from a strict Poisson distribution with parameter 5. The variability of embryo yield was therefore lower than in *Model 2* ($CV = 0.45$, $R = 0.00$).

Model 4

Finally, a model in which a constant number of embryos collected per flush and per donor (5 embryos) was considered ($CV = R = 0.00$).

Figure 1 shows the distribution of the number of transferable embryos under the first 3 models.

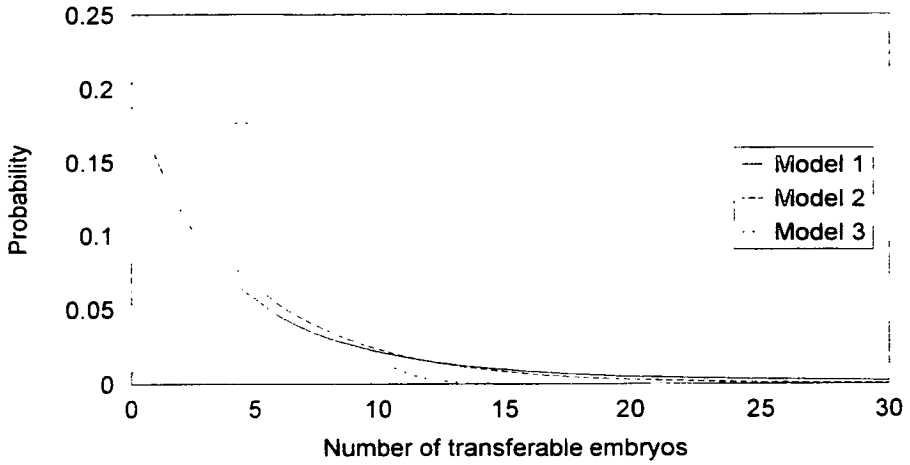


Fig 1. The distribution of the number of transferable embryos per flush and per donor under different models for embryo yield.

Comparison among breeding schemes

The scheme with current values for reproductive parameters was used as a point of reference for comparisons. Average true breeding values (G_i) and inbreeding coefficients (F_i) of individuals born at the ith year were obtained. Annual rate of response between years j and i was calculated as $\Delta G_{i-j} = (G_j - G_i)/(j-i)$, where $j > i$. Rates of inbreeding were obtained every year as $\Delta F_i = (F_i - F_{i-1})/(1-F_{i-1})$. The rate of inbreeding between years j and i (ΔF_{i-j}) was obtained by taking the average of annual rates. Also, the following parameters were calculated in the simulations: 1) genetic variance of animals born every year; 2) accuracy of selection (correlation between the true breeding values and selection criteria of the candidates for selection); 3) genetic selection differentials (difference between the mean values of selection criteria of selected individuals and candidates for selection) and selection intensities for males and females; 4) generation intervals (average age of parents when offspring are born) for males and females; and 5) variance of family sizes for male and female parents. The latter was calculated as described in Villanueva *et al* (1994). Each scheme was replicated 200 times and the values presented are the average over all replicates. The criteria for comparing different schemes were the rates of response (ΔG_{15-25}) and inbreeding (ΔF_{15-25}) at the later years (from year 15 to year 25). The cumulative response (G_{25}) and inbreeding at year 25 (F_{25}) were also compared.

RESULTS

Models for embryo production

Table I shows the genetic progress and the inbreeding obtained under different models used to generate the number of embryos per collection. The results show that the inbreeding obtained depended on the values of the coefficient of variation and the repeatability of embryo yield. By making the correlation between embryo production at different recoveries equal to zero ($R = 0.00$), the rate of inbreeding decreased by 4% (*Models 1* and *2*). The effect of the coefficient of variation of embryo yield on the rate of inbreeding was notable. By increasing the coefficient of variation by a factor of 2.4 the rate of inbreeding increased by 14% (*Models 2* and *3*). The rate of inbreeding obtained when the number of embryos collected was fixed (*Model 4*) was between 2 and 14% lower than that obtained when there was variability in embryo yield but the repeatability was zero (*Models 3* and *2*). The genetic progress decreased as variability of embryo production increased. The decrease in response was however small. The genetic gain obtained with *Model 1* was around 2% lower than that obtained with *Model 4* (fixed number of embryos).

Table I. Genetic progress (phenotypic standard deviation units) and inbreeding (%) and their standard errors using different models to generate the number of embryos collected per flush and per donor[a].

Model[b]	*Genetic progress*		*Inbreeding*	
	$\Delta G_{15\text{-}25}$	G_{25}	$\Delta F_{15\text{-}25}$	F_{25}
Model 1	0.199 (0.002)	4.85 (0.03)	1.98 (0.07)	40.54 (0.53)
Model 2	0.203 (0.002)	4.89 (0.02)	1.90 (0.06)	38.61 (0.55)
Model 3	0.201 (0.001)	4.92 (0.03)	1.67 (0.05)	35.72 (0.45)
Model 4	0.204 (0.002)	4.92 (0.03)	1.64 (0.05)	35.23 (0.50)

[a]Standard errors are shown in brackets; [b]*Model 1*: Poisson with variable parameter ($MEAN = 5.0$, $CV = 1.28$, $R = 0.22$). *Model 2*: Poisson with variable parameter ($MEAN = 5.0$, $CV = 1.09$, $R = 0.00$). *Model 3*: Poisson with constant parameter ($MEAN = 5.0$, $CV = 0.45$, $R = 0.00$). *Model 4*: fixed number of embryos ($MEAN = 5.0$, $CV = 0.00$, $R = 0.00$).

Improved embryo recovery and embryo transfer

Values for reproductive parameters utilized in different schemes are shown in table II. Two different situations of improved technology for embryo production were considered. Firstly, the coefficient of variation of embryo yield was decreased and the mean was maintained. Secondly, the coefficient of variation was decreased and the mean was increased. Under *Model 1*, the coefficient of variation can be decreased by increasing v since $CV = [(1 + \beta_i^2)/\beta_i v]^{1/2}$. In order to keep the mean constant, β_i must be decreased, which is achieved by decreasing μ. In the second situation (coefficient of variation decreased and mean increased) the parameter v

Table II. Values for reproductive parameters used in the simulations.

Parameter	*Current value*	*Future values*
Mean number of transferable embryos per flush and donor (*MEAN*)	5.0	7.4, 9.6
Coefficient of variation of embryo production (*CV*)	1.28	1.10, 1.07, 1.05, 0.91
Repeatability of embryo production (*R*)	0.22	0.34, 0.38, 0.29, 0.35
Frequency of flushing in days between flushes (*FC*)	60	45
Embryo survival rate (*ESR*)	0.55	0.65, 0.75

Table III. Values of embryo distribution parameters used in the simulations and expected mean (*MEAN*), coefficient of variation (*CV*) and repeatability (*R*) of embryo yield.

Distribution parameter			*Embryo yield parameter*		
ν	μ	σ^2	MEAN	CV	R
1.0	1.46	0.4	5.0	1.28	0.22
2.0	0.77	0.4	5.0	1.10	0.34
2.5	0.50	0.4	5.0	1.07	0.38
1.5	1.46	0.4	7.4	1.05	0.29
2.0	1.46	0.4	9.6	0.91	0.35

was increased whereas μ was kept constant. Values used for embryo distribution parameters as well as the resulting *MEAN*, *CV* and *R* are shown in table III.

The rates of response and inbreeding obtained with improved values for parameters of embryo recovery and embryo transfer are shown in table IV. The first row of the table represents the current situation and is used as a reference. The expected number of embryos transferred for each scheme is shown in the last column. Decreasing the coefficient of variation of embryo production while keeping the mean approximately constant did not have an effect on rates of response and inbreeding. This may be due to the increased repeatability that accompanied the decrease in *CV* in the model. The influence of the repeatability of embryo yield has been shown in the previous section. Increasing the mean number of embryos transferred led to a notable increase in the rates of response, due to increased selection intensities and accuracy and decreased generation intervals. In this case, the number of calves born per year (N_{CB}) was unrestricted and the number of donors was constant, so increasing embryo yield led to more candidates for selection. Male and female selection intensities (i_{σ} and i_{ς}) and generation intervals (L_{σ} and L_{ς}) are shown in table V. The rate of inbreeding (per year and per generation) was also increased (particularly when the mean number of embryos produced was 9.6) due to increased full-sib family sizes and intensities of selection and decreased generation intervals.

The assumed current frequency of collection of embryos (*FC*) was 60 d (3 flushes in a 6 month period). The potential benefits from increasing the frequency of

Table IV. Effect of embryo recovery and embryo transfer parameters on genetic progress (phenotypic standard deviation units) and inbreeding (%) when the number of calves born per year is unconstrained[a].

ESR[b]	FC	MEAN	CV	R	*Genetic progress*		*Inbreeding*		N_{CB}	E(ET)
					$\Delta G_{15\text{-}25}$	G_{25}	$\Delta F_{15\text{-}25}$	F_{25}		
0.55	60	5.0	1.28	0.22	0.199 (0.002)	4.85 (0.03)	1.98 (0.07)	40.54 (0.53)	297.8 (1.4)	540
		5.0	1.10	0.34	0.204 (0.002)	4.93 (0.03)	2.02 (0.07)	40.80 (0.57)	306.1 (1.6)	540
		5.0	1.07	0.38	0.200 (0.002)	4.84 (0.03)	1.93 (0.06)	39.91 (0.52)	293.1 (1.4)	540
		7.4	1.05	0.29	0.222 (0.001)	5.33 (0.03)	2.08 (0.08)	43.86 (0.61)	440.2 (1.8)	799
		9.6	0.91	0.35	0.236 (0.002)	5.67 (0.03)	2.17 (0.08)	45.34 (0.59)	571.1 (2.1)	1 037
	45	5.0	1.28	0.22	0.216 (0.002)	5.23 (0.02)	2.05 (0.07)	42.98 (0.59)	400.0 (1.8)	720
0.65	60	5.0	1.28	0.22	0.212 (0.002)	5.04 (0.03)	2.15 (0.08)	41.46 (0.61)	354.6 (1.7)	540
0.75	60	5.0	1.28	0.22	0.216 (0.002)	5.23 (0.02)	2.08 (0.08)	43.34 (0.64)	413.0 (1.8)	540

[a]Standard errors are shown in brackets; [b]*ESR* = embryo survival rate, *FC* = frequency of embryo collection (days between flushes), N_{CB} = number of calves born per year, *E(ET)* = expected number of embryos transferred.

Table V. Effect of embryo recovery and embryo transfer parameters on selection intensities (i_{σ} and i_{φ}) and generation intervals (L_{σ} and L_{φ}) when the number of calves born per year is unconstrained [a].

ESR[b]	FC	MEAN	CV	R	i_{σ}	i_{φ}	L_{σ}	L_{φ}
0.55	60	5.0	1.28	0.22	1.62 (0.005)	1.35 (0.004)	2.95 (0.007)	3.15 (0.006)
		5.0	1.10	0.34	1.64 (0.005)	1.37 (0.004)	2.94 (0.007)	3.14 (0.006)
		5.0	1.07	0.38	1.62 (0.005)	1.35 (0.004)	2.96 (0.007)	3.15 (0.006)
		7.4	1.05	0.29	1.80 (0.005)	1.52 (0.004)	2.84 (0.006)	3.03 (0.005)
		9.6	0.91	0.35	1.91 (0.004)	1.64 (0.004)	2.80 (0.005)	2.97 (0.004)
	45	5.0	1.28	0.22	1.76 (0.005)	1.48 (0.004)	2.87 (0.007)	3.05 (0.006)
0.65	60	5.0	1.28	0.22	1.70 (0.005)	1.43 (0.004)	2.89 (0.005)	3.09 (0.005)
0.75	60	5.0	1.28	0.22	1.76 (0.005)	1.49 (0.004)	2.86 (0.006)	3.05 (0.006)

[a]Standard errors are shown in brackets. [b]ESR = embryo survival rate, FC = frequency of embryo collection (days between flushes).

flushing to 45 d (4 flushes in a 6 month period) on rates of response and inbreeding are also shown in table IV. The increase in flushing frequency to this optimistic future value produced a clear increase in genetic progress. This increase in genetic progress was due to increased selection intensities and accuracy of selection and decreased generation intervals (table V). Inbreeding was slightly higher when donors were flushed 4 times per period. Finally, by increasing the probability that an embryo survives until calving (ESR) from 0.55 to 0.65 and 0.75, cumulative genetic response was increased by 4 and 8%, respectively. The rate of inbreeding was also increased. Table V indicates increases in selection intensities and decreases in generation intervals with improved viability of the embryos.

Tables IV and V show results obtained without a constraint on the number of calves born in the scheme. By increasing the mean number of embryos per flush and per donor, the frequency of flushing or the embryo survival probability, the expected number of offspring is increased. Genetic progress obtained at year 25 was directly proportional to the number of calves born each year (table IV). With more offspring born, the selection intensities and the accuracy of selection were increased and the generation intervals were decreased (table V). Also, the rate of inbreeding (per year and per generation) was increased by improving embryo transfer and embryo recovery techniques. For a fixed number of sires and donors, the increase in the number of offspring born per year led to an increase in the variances of family sizes.

Comparing schemes which differ widely in the number of offspring produced is unfair. This is because genetic gains are expected to be higher (and inbreeding is expected to be lower) in larger schemes, irrespective of the use of breeding technologies. Also, in practice, there will usually be a limitation on the number of embryo collections or transfers which can be made, the number of recipients available, or the number of testing places available for calves. These constraints are equivalent, except when different success rates are assumed for embryo technologies.

Simulations were therefore also run with a restriction on the number of offspring born every year and the results are presented in table VI.

When the mean number of embryos collected per flush and the frequency of collection were increased, some embryos were discarded in order to transfer a fixed number. In these cases, the expected average number of embryo transfers per year was 540 (this is the expected number of transfers with current values for reproductive parameters and 18 donors). Embryos were not discarded at random; most were discarded from donors producing more embryos, in order to equalize family sizes. The decision on which embryos were transferred was made within individual flushes. First, the number of embryos recovered from each donor was obtained using *Model 1*. After that, 1 embryo from each donor (if available) was allocated (for transfer) in succession and this process was repeated until the desired total number of embryos was reached. In these cases, the maximum number of embryos transferred after a single flush was $18 \times 5 = 90$. When the survival rate of embryos from transfer until birth was increased, the number of transfers was decreased in order to obtain a fixed number of calves born per year. Again, more embryos were discarded from donors with higher embryo production. Table VI shows that with these strategies, the number of calves born per year (N_{CB}) was approximately constant in all schemes.

Increasing the mean number of embryos produced per flush and per donor from 5.0 to 7.4 and 9.6 decreased the rate of inbreeding by up to 10% even with the increased repeatability (table VI). Thus, restricting family sizes nullified the effect of repeatability. The decrease in inbreeding rates was due to decreased variances of family sizes. The increase in frequency of flushing also led to a decrease in the rate of inbreeding (by 11%) whereas the genetic progress was not affected. Finally, by increasing the probability of embryo survival, the rate of inbreeding was reduced by up to 5% with no effect on response. These latter schemes ($ESR > 0.55$) were not compared on an equal basis with the others, since fewer embryos were transferred and less recipients were used.

Sexing of semen

Table VII shows the results obtained when sexed semen was used to change the sex ratio from 1:1 to 1:2 and 1:3 in favour of females, in order to increase the selection intensity in this sex. The number of embryos obtained per flush and per donor was simulated using *Model 1*. The number of transfers per year was expected to be the same for all schemes (540 transfers per year). There was no benefit from using sexed semen when the number of progeny tested per year was fixed. Table VIII shows generation intervals and selection intensities for schemes with different sex ratios. As expected, the selection intensity of females increases as the proportion of female offspring increases. However, at the same time, there is a reduction in the selection intensity of males and the average selection intensity is not increased. This led to similar rates of response when different sex ratios were simulated. Generation intervals were very similar for schemes with different sex ratios.

Table VI. Effect of embryo recovery and embryo transfer parameters on genetic progress (phenotypic standard deviation units) and inbreeding (%) when the number of calves born per year is constrained[a].

ESR[b]	FC	MEAN	CV	R	*Genetic progress*		*Inbreeding*		N_{CB}	E(ER)	E(ET)
					$\Delta G_{15\text{-}25}$	G_{25}	$\Delta F_{15\text{-}25}$	F_{25}			
0.55	60	5.0	1.28	0.22	0.199 (0.002)	4.85 (0.03)	1.98 (0.07)	40.54 (0.53)	297.8 (1.4)	540	540
		7.4	1.05	0.29	0.205 (0.002)	4.95 (0.03)	1.85 (0.06)	38.21 (0.54)	299.8 (0.3)	799	540
		9.6	0.91	0.35	0.204 (0.002)	4.92 (0.03)	1.79 (0.06)	36.24 (0.53)	303.1 (0.2)	1 037	540
	45	5.0	1.28	0.22	0.203 (0.002)	4.87 (0.03)	1.77 (0.06)	38.10 (0.51)	288.7 (0.5)	720	540
0.65	60	5.0	1.28	0.22	0.201 (0.002)	4.84 (0.03)	1.93 (0.06)	39.44 (0.54)	283.8 (0.6)	540	457
0.75	60	5.0	1.28	0.22	0.202 (0.002)	4.85 (0.03)	1.89 (0.06)	38.60 (0.55)	291.4 (0.4)	540	396

[a]Standard errors are shown in brackets. [b]*ESR* = embryo survival rate, *FC* = frequency of embryo collection (days between flushes), N_{CB} = number of calves born per year, *E(ER)* = expected number of embryos recovered, *E(ET)* = expected number of embryos transferred.

Table VII. Effect of sexing of semen on genetic progress (phenotypic standard deviation units) and inbreeding (%) [a].

Sex ratio (♂/♀)	*Genetic progress*		*Inbreeding*	
	$\Delta G_{15\text{-}25}$	G_{25}	$\Delta F_{15\text{-}25}$	F_{25}
1:1	0.199 (0.002)	4.85 (0.03)	1.98 (0.07)	40.54 (0.53)
1:2	0.200 (0.002)	4.83 (0.03)	2.09 (0.07)	41.90 (0.55)
1:3	0.195 (0.002)	4.70 (0.03)	2.03 (0.07)	40.79 (0.61)

[a]Standard errors are shown in brackets.

Table VIII. Effect of sexing of semen on selection intensities ($i_♀$ and $i_♀$) and generation intervals ($L_♂$ and $L_♀$) [a].

Sex ratio (♂/♀)	$i_♂$	$i_♀$	$L_♂$	$L_♀$
1:1	1.62 (0.005)	1.35 (0.004)	2.95 (0.007)	3.15 (0.006)
1:2	1.44 (0.004)	1.49 (0.004)	2.99 (0.007)	3.11 (0.006)
1:3	1.30 (0.004)	1.54 (0.004)	3.04 (0.007)	3.13 (0.007)

[a]Standard errors are shown in brackets.

DISCUSSION

Two novelties of the present study are that the coefficient of variation of embryo yield used corresponds to that observed in real data (and is higher than that used in previous studies) and that the repeatability of embryo yield has been included. Studies evaluating the use of MOET for genetic improvement of ruminants have frequently assumed a fixed number of embryos per collection (Land and Hill, 1975; Nicholas and Smith, 1983; Juga and Mäki-Tanila, 1987; Gearheart *et al*, 1989; Keller *et al*, 1990; Meuwissen, 1991; Ruane and Thompson, 1991; Toro *et al*, 1991; Bondoc and Smith, 1993; Leitch *et al*, 1994). Several studies have considered variable family sizes but using hypothetical distributions.

Ruane (1991) simulated variable family sizes by obtaining the number of embryos recovered per donor from a normal distribution with mean 16 and variance up to 64 (he assumed 4 flushes per generation and an average of 4 embryos per flush). Thus the highest coefficient of variation for embryo production simulated was 0.50. His results showed a small effect of variation in family sizes on rates of response and inbreeding. The reduction in response by including variation in the number of embryos collected per donor was around 4% whereas the increase in inbreeding was up to 3%. Colleau (1991) used a 'scaled' binomial distribution to model embryo production. The proportion of treated cows responding to superovulation (p) was 0.7 and the number of embryos collected per flush and per treated cow (m) was 4 or 5. Thus the number of embryos collected per flush and per donor was $m\theta$, where θ is a random variable having a binomial distribution with the parameters $n = 1$ (flushes) and $p = 0.7$ (probability of a successful flush). The mean of this

distribution is mnp and the variance is $m^2np(1-p)$ which leads to a coefficient of variation of $[(1-p)/p]^{1/2} = 0.65$. Poisson distributions (with constant parameters) have been used also for modelling the number of embryos recovered following superovulation. Schrooten and van Arendonk (1992) used a Poisson distribution with parameter 5 to obtain the number of live calves per flush. This implies a coefficient of variation of 0.45.

More realistic models have been proposed to generate embryo yield distributions (Foulley and Im, 1993; Tempelman and Gianola, 1993; 1994). In a simulation study, Tempelman and Gianola (1994) generated embryo yields in MOET nucleus herds using a model which includes repeatable variation among donors. Within-dam variation was modelled using a Poisson distribution. However, Woolliams *et al* (1995) have shown that extra-Poisson variation is observed in practice. Actual data on ovulation rates and embryo recoveries are better described by negative binomial distributions than by Poisson models. In the model of Tempelman and Gianola (1994), more within-donor variation could be generated by including an additional random term.

Wray and Simm (1990) used a distribution based on commercial data to generate the number of calves per flush in beef cattle. The coefficient of variation used was 1.15. The coefficient of variation of the number of calves born per flush can be obtained from the mean and the coefficient of variation of embryo yield as

$$[ESR(1-ESR)MEAN + ESR^2CV^2MEAN^2]^{1/2}/[ESR \times MEAN]$$

The value used in the present study (1.34) is higher than that used in previous studies. Changes in inbreeding also depend on the value of repeatability of embryo yield (table I). With increased repeatability of embryo yield, an increase in the variance of family size is expected (since the variance between donors increases), with the potential for fewer families to make a greater contributions to successive generations. Thus, models used previously (which have assumed a constant number of embryos per collection or a variable number, but with lower coefficient of variation and no correlation between different recoveries) have underestimated the rate of inbreeding and overestimated genetic gain.

Improved embryo recovery and transfer success rates lead to higher rates of response and inbreeding than current success rates, providing the number of calves tested per year is unconstrained. This is due to higher selection intensities and accuracies, lower generation intervals and higher and more variable family sizes. However, it is unrealistic to assume unconstrained number of calves born since, in centralized nucleus herds, costs will depend of the total number of animals in the scheme. When different schemes are compared at a constant number of offspring born per year, improved success rates do not increase progress, since selection intensities and generation intervals are maintained. However, inbreeding rates can be markedly reduced by discarding more embryos from donors with the highest embryo production, to equalize family sizes. Although the results presented are for the later years of selection, rates of response and inbreeding were also obtained for the early years (from year 5 to year 15). There was no significant effect of time on the results of comparisons among schemes.

Schemes which assumed improved embryo transfer techniques (improved ESR) require less recipients, and therefore, should have lower costs than the rest of the

schemes (see $E(ET)$ in table VI). An alternative basis for making fair comparisons would be to transfer 540 embryos and, from the 351 ($ESR = 0.65$) and 405 ($ESR = 0.75$) calves expected, choose for performance testing the 297 animals which would best equalize family sizes. In this case, schemes with improved ESR could benefit from selling surplus calves. On the other hand, improved embryo recovery techniques (increased $MEAN$ and FC) give the opportunity of selling surplus embryos (see $E(ER)$ in table VI). However, this benefit is difficult to quantify as there is not currently a large or stable market for beef cattle embryos.

Variation in response to selection can be an important limitation of MOET nucleus schemes. Nicholas (1989) suggested that the maximum acceptable coefficient of variation of response after 10 years of selection ranged from 5 to 10%. For the schemes considered in this paper, the coefficients of variation of response over a 10 year period varied from 10 to 14%. Thus, the size of the nucleus needs to be larger than that considered here, or strategies for controlling inbreeding should be applied, to have a reasonable level of risk.

In breeding programmes, sexing of embryos or semen could be used to increase the selection intensity applied to females. However, there seems to be no benefit from sexing when performance information is available on both sexes and comparisons are made at a fixed number of individuals tested per year (table VII). As expected, the selection intensity of females increased as the proportion of female offspring increased. However, at the same time, there was a reduction in the selection intensity of males and the overall selection intensity was unchanged (table VIII). Wray and Goddard (1994) have investigated a different strategy, in sheep MOET schemes, which used sexed semen (female) to inseminate the selected dams with lower EBVs whereas the top selected dams were inseminated with unsexed semen (to avoid decreases in male selection intensity). However, in MOET schemes, the overall selection intensity did not change. They found a benefit from sexing in conventional schemes (without MOET), suggesting that sexing can be beneficial when the male and female selection intensities differ greatly.

On the other hand, Colleau (1991) reported, for dairy cattle, slightly higher gains for adult MOET schemes with sexing of embryos than for juvenile MOET schemes without sexing. Since juvenile schemes are expected to be superior to adult schemes without sexing (with respect to genetic progress), his results suggest a clear benefit from sexing. Whereas the studies discussed in the previous paragraph have considered fixed numbers of transfers, sires and dams, Colleau's model allowed the number of dams to vary. The nucleus considered assumed females dispersed across many recorded herds. For his adult scheme, the overall number of dams used for replacements was much higher than in the juvenile scheme. This allowed selection differentials in the adult scheme to be high enough to compensate for the longer generation intervals. In centralized nucleus schemes, a constraint on the number of dams would be also needed. In these circumstances the advantage of sexing would be doubtful.

In conclusion, the values of the coefficient of variation and the repeatability of embryo yield are important in determining rates of inbreeding. When the number of testing places is constrained, improved technologies can greatly decrease the rate of inbreeding without affecting genetic gain. Finally, when performance information

is available on both sexes and comparisons are carried out at a fixed number of individuals tested per year, there is no apparent benefit in response from sexing.

ACKNOWLEDGMENTS

We would like to thank Drs B McGuirk, R Thompson and a referee for their valuable suggestions. This work was funded by the Ministry of Agriculture, Fisheries and Food, the Milk Marketing Board of England and Wales and the Meat and Livestock Commission. SAC receives financial support from the Scottish Office Agriculture and Fisheries Department and the Roslin Institute receives financial support from the Biotechnology and Biological Sciences Research Council and the Ministry of Agriculture, Fisheries and Food.

REFERENCES

Bondoc OL, Smith C (1993) Optimized testing schemes using nucleus progeny, adult MOET siblings, or juvenile MOET pedigrees in dairy cattle closed populations. *J Anim Breed Genet* 101, 30-40

Colleau JJ (1991) Using embryo sexing within closed mixed multiple ovulation and embryo transfer schemes for selection of dairy cattle. *J Dairy Sci* 74, 3973-3984

Dekkers JCM (1992) Structure of breeding programs to capitalize on reproductive technology for genetic improvement. *J Dairy Sci* 75, 2880-2891

Foulley JL, Im S (1993) A marginal quasi-likelihood approach to the analysis of Poisson variables with generalized linear mixed models. *Genet Sel Evol* 25, 101-107

Gearheart WW, Smith C, Teepker G (1989) Multiple ovulation and embryo manipulation in the improvement of beef cattle: relative theoretical rates of genetic change. *J Anim Sci* 67, 2863-2871

Juga J, Mäki-Tanila (1987) Genetic change in a nucleus breeding dairy herd using embryo transfer. *Acta Agric Scand* 37, 511-519

Keller DS, Gearheart WW, Smith C (1990) A comparison of factors reducing selection response in closed nucleus breeding schemes. *J Anim Sci* 68, 1553-1561

Land RB, Hill WG (1975) The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Anim Prod* 21, 1-12

Leitch HW, Smith C, Burnside EB, Quinton M (1994) Genetic response and inbreeding with different selection methods and mating designs for nucleus breeding programs of dairy cattle. *J Dairy Sci* 77, 1702-1718

Lohuis MM, Smith C, Dekkers JCM (1993) MOET results from a dispersed hybrid nucleus programme in dairy cattle. *Anim Prod* 57, 369-378

Luo ZW, Woolliams JA, Simm G (1994) An assessment of present and future effectiveness of embryological techniques in ruminants. *Agric Biotech News Infor* 6, 13N-18N

Meuwissen THE (1991) The use of increased female reproductive rates in dairy cattle breeding schemes. *Anim Prod* 52, 21-31

Nicholas FW (1989) Incorporation of new reproductive technology in genetic improvement programmes. *In: Evolution and Animal Breeding* (WG Hill, FC MacKay, eds), CBA International, Wallingford, UK, 203-209

Nicholas FW, Smith C (1983) Increased rates of genetic changes in dairy cattle by embryo transfer and splitting. *Anim Prod* 36, 341-353

Ruane J (1991) The importance of family sizes in adult multiple ovulation and embryo transfer (MOET) nucleus breeding schemes in dairy cattle. *Anim Prod* 52, 33-47

Ruane T, Thompson R (1991) Comparison of simulated and theoretical results in adult MOET nucleus schemes for dairy cattle. *Livest Prod Sci* 28, 1-20

Schrooten C, van Arendonk JAM (1992) Stochastic simulation of dairy cattle breeding schemes: genetic evaluation of nucleus size and type. *J Anim Breed Genet* 109, 1-15

Smith C (1986) Use of embryo transfer in genetic improvement of sheep. *Anim Prod* 42, 81-88

Tempelman RJ, Gianola D (1993) Marginal maximum likelihood estimation of variance components in Poisson mixed models using Laplacian integration. *Genet Sel Evol* 25, 305-319

Tempelman RJ, Gianola D (1994) Assessment of a Poisson animal model for embryo yield in a simulated multiple ovulation-embryo transfer scheme. *Genet Sel Evol* 26, 263-290

Toro M, Silió L, Perez-Enciso M (1991) A note on the use of mate selection in closed MOET breeding schemes. *Anim Prod* 53, 403-406

Villanueva B, Woolliams JA, Simm G (1994) Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genet Sel Evol* 26, 517-535

Woolliams JA, Luo ZW, Villanueva B *et al* (1995) Analysis of factors affecting superovulatory responses in ruminants. *J Agric Sci, Camb* 124, 61-70

Way NR, Goddard ME (1994) MOET breeding schemes for wool sheep. I. Design alternatives. *Anim Prod* 59, 71-86

Wray NR, Simm G (1990) The use of embryo transfer to accelerate genetic improvement in beef cattle. *In: Proceedings of the 4th World Congress on Genetics Applied to Livestock Production, Vol XV, Edinburgh, 23-27 July* (WG Hill, R Thompson, JA Woolliams, eds), 315-318

Paper 22

Animal Science 1995, **61**: 231-239

1357-7298/95/45610231$20·00

Genetic progress and inbreeding for alternative nucleus breeding schemes for beef cattle

B. Villanueva[1], G. Simm[1] and J. A. Woolliams[2]

[1]*Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG*
[2]*Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS*

Abstract

Alternative closed breeding schemes for beef cattle are analysed using stochastic computer simulation. Multiple ovulation and embryo transfer (MOET) schemes are compared with conventional schemes (schemes without MOET) with an equal expected number of progeny born per year. Schemes are compared for genetic gain and inbreeding obtained after 25 years of selection. The trait considered, evaluated in both sexes, has an initial heritability of 0·35. Different population sizes and numbers of sires selected are evaluated. Current realistic parameters for embryo production are assumed in MOET schemes.

After 25 years of selection, and with no control on inbreeding, cumulative genetic gains are about 50% higher in MOET schemes compared with conventional schemes. The benefit from MOET is mostly due to increased selection intensities in females. The rate of inbreeding increases by up to nearly 300% when MOET is used. This maximum percentage increase in inbreeding following the use of MOET can be reduced to about 100% when selection and mating strategies for controlling inbreeding are used. The effect of the number of sires used on the inbreeding obtained is more important than the effect of the size of the herd. In MOET schemes, increasing the number of sires selected by a factor of three, leads to reductions in inbreeding rates of 40%. When schemes of the same size are compared at similar acceptable inbreeding levels, MOET schemes give around 30% higher genetic progress than conventional schemes.

Keywords: *beef cattle, breeding programmes, inbreeding, MOET, selection responses.*

Introduction

There are several biological and structural limitations to accelerating genetic improvement of beef cattle. Cattle reach sexual maturity at a later age, and have a lower reproductive rate than pigs, poultry, and to a lesser extent sheep. This imposes a biological limit on the generation intervals and selection intensities which can be achieved with natural mating. However, these limits can be raised by the appropriate use of techniques such as AI (artificial insemination) and MOET (multiple ovulation and embryo transfer), with corresponding improvements in the rate of genetic gain. Often there are also structural limitations on the rate of genetic improvement of beef cattle. Foremost amongst these is the small size of pedigree herds in many countries, which limits the intensity and accuracy of selection. In most European countries, the average pedigree herd size is smaller than 35 cows for most breeds (Simm, Steane and Wray, 1990). This is exacerbated if the method of evaluation, or a lack of genetic links across herds, does not permit comparison of EBVs (estimated breeding values) across herds. In theory, large nucleus breeding schemes could overcome some of these limitations and, perhaps particularly if combined with MOET, could allow faster genetic improvement.

Early studies on the use of MOET in cattle nucleus breeding schemes (Land and Hill, 1975; Nicholas, 1979; Nicholas and Smith, 1983) overestimated the expected extra response to selection, and substantially underestimated the expected extra inbreeding. These predictions assumed discrete generations and constant family size. Also, they did not account for reduction in genetic variance due to selection and inbreeding nor for reduction in selection differentials due to correlations among EBVs of relatives. Subsequently there has been a great deal of research effort, particularly in dairy cattle, to refine these expectations and to investigate

optimum designs of MOET breeding schemes (see review of Villanueva and Simm, 1994).

Wray and Simm (1990) showed that when schemes for beef cattle are compared at the same rate of inbreeding then the use of MOET can result in up to 50% extra gain. However, they still assumed that there was no correlation between embryo production in successive collections and assumed a coefficient of variation lower than that observed in practice. The coefficient of variation and the repeatability of embryo yield have been shown to be important parameters in determining the rate of inbreeding (Villanueva, Woolliams and Simm, 1995).

The use of MOET can lead to proportionally higher increases in rates of inbreeding than in rates of genetic gain (e.g. Wray and Simm, 1990). However, there are selection and mating procedures which can substantially reduce the rate of inbreeding with a minimal effect on genetic gain. Using stochastic simulation to model MOET schemes in beef cattle, Villanueva, Woolliams and Simm (1994) examined the effect of combining factorial mating designs (Woolliams, 1989) with selection on BLUP (best linear unbiased prediction) EBVs using a deliberately increased heritability (Grundy, Caballero, Santiago and Hill, 1994). They showed that, by using these strategies, the rate of inbreeding can be decreased by over 50% with only a very small reduction in response (with respect to selection on standard BLUP EBVs and nested mating designs).

Although there are several dairy MOET nucleus breeding schemes in operation worldwide, there are relatively few examples (e.g. Broadbent, 1990; Scott, 1992) of the structured use of MOET in beef cattle breeding. This is at least partly due to the very high levels of financial investment required to establish and operate such schemes (and perhaps also due to the additional complexity of managing breeding schemes using MOET). If more of these schemes are to be established in future, it is important for potential investors that there are accurate expectations of the extra benefits, and that practical guidelines are produced to allow these benefits to be achieved. The purpose of this study is to reappraise the potential extra genetic gains, and the potential extra inbreeding, in beef cattle MOET schemes, after accounting for what are known now to be deficiencies in some of the early studies. This is vitally important for people intending to invest in MOET schemes and to date no one has given a clear statement of whether such schemes are still genetically appropriate in the beef industry.

In this study the methods for reducing inbreeding described above were applied and results were obtained with more realistic models than those used in the past. A range of mating ratios and nucleus herd sizes were examined, and the effectiveness of alternative breeding schemes was assessed by comparing genetic gains and rates of inbreeding obtained after 25 years of selection.

Methods

Closed nucleus schemes with overlapping generations were generated for beef cattle using stochastic simulation. The trait under selection had an initial heritability of 0·35 and it was recorded in both sexes at around 400 days of age (between 385 and 415 days), at the end of a performance test. This trait can be interpreted as live weight, daily gain or some *in vivo* estimates of carcass characteristics. An additive infinitesimal genetic model was assumed. Different numbers of males (N_S) and females (N_D) constituting the base population were considered. Base population animals were assumed to be unrelated. Their true breeding values were obtained from a normal distribution with mean zero and variance (σ^2_A) 0·35. Phenotypic values were obtained by adding an environmental component taken at random from a normal distribution with mean zero and variance 0·65. Equal numbers of animals aged 2, 3 and 4 years were generated.

Selection was carried out for 25 years. Animals were genetically evaluated twice every year (evaluation period = 6 months). Selection took place also twice every year. EBVs were obtained using two different procedures: (1) individual animal model BLUP using the true heritability (standard BLUP); and (2) individual animal model BLUP using an artificially raised heritability, h^2_{AR} = 0·7 (modified BLUP). Villanueva *et al.* (1994) showed that this value for h^2_{AR} is the optimum in the sense that inbreeding can be substantially reduced without affecting response. Higher values for h^2_{AR} decreased also genetic gain. The number of breeding males and females was constant in each evaluation period and equal to the number of base males and females (N_S sires and N_D dams). Males and females with the highest EBVs were selected and randomly mated under a nested or a factorial design. Under the nested design, donor females were mated to the same sire in successive flushes within an evaluation period, whereas under the factorial design, donor females were mated to different sires in successive flushes. The number of matings per sire was approximately equal for each sire in one evaluation period. Animals were selected irrespective of whether they had been selected in previous periods and animals not selected were culled from the herd. No animals were culled before being performance tested. There were no restrictions

on the number of selected animals per family. The minimum age for breeding for females was 15 months. Males were assumed to be capable of breeding after being performance tested. Breeding took place throughout the year.

True breeding values of the offspring were generated as:

$$TBV_I = (1/2)\,(TBV_S + TBV_D) + m_I$$

where TBV_I, TBV_S and TBV_D are the true breeding values of the individual I, its sire and its dam, respectively, and m_I is the Mendelian sampling term. The latter was taken at random from a normal distribution with mean zero and variance $(1/2)\,[1 - (F_S + F_D)/2]\sigma_A^2$, where F_S and F_D are the inbreeding coefficients of the sire and dam, respectively. Inbreeding coefficients were obtained from the additive relationship matrix as described by Quaas (1976). The sex of the offspring was assigned at random with equal probability. All animals in the population had one of five different status codes (unborn, immature, mature, breeding and culled or dead) at any time. The status codes of the animals were checked every day and changed at key times as appropriate (except for the status 'culled or dead' which is definitive). At all ages after birth, individuals were subject to a mortality rate that varied with age. Mortality rates from birth to 3 weeks, 6 months and 2, 5, 10 or 15 years were 2%, 3%, 4%, 7%, 14% and 100%, respectively. No animal was assumed to live beyond 15 years.

Conventional schemes

The method of genetic evaluation used in conventional schemes was standard BLUP. Selected cows that were in oestrus were naturally mated to selected sires under a nested mating design. A natural mating had a 63% chance of success. Cows were given three opportunities to become pregnant. The very first oestrus took place at 365 + ø days of age, where ø is a random normal deviate sampled from a normal distribution with mean zero and variance 25. Thus, the first calving was expected at 2 years of age. The first oestrus following a calving was assumed to occur 60 days after that calving. The oestrous cycle was assumed to be 21 days. The calving date was simulated as $DM + 280 + \gamma$, where DM is the day of mating and γ is a random normal deviate sampled from a normal distribution with mean zero and variance 25. The term γ represents variation in gestation length. Bourdon and Brinks (1982) gave an estimate of around five for the phenotypic standard deviation of gestation length in beef cattle. Gestation lengths may be longer than assumed here in some breeds. Also, the expected age at first calving may be longer than 2 years. However, this should not affect comparisons among schemes since the same figures were used for MOET schemes. Mortality rate of embryos from conception to calving was 6%.

MOET schemes

Simulation of MOET schemes has been described in detail by Villanueva *et al.* (1994). Values used for embryo recovery and embryo transfer represent current values for these reproductive technologies (Luo, Woolliams and Simm 1994; Woolliams, Luo, Villanueva, Waddington, Broadbent, McKelvey and Robinson, 1994). The number of transferable embryos collected per flush and cow was obtained from a Poisson distribution whose parameter is distributed according to a gamma distribution (Woolliams *et al.*, 1994, Villanueva *et al.*, 1995). This model describes very accurately embryo yield in practice (Woolliams *et al.*, 1994). The mean number of transferable embryos per flush and per donor was 5·1 and the coefficient of variation and repeatability of embryo yield were 1·28 and 0·22, respectively. These values were obtained from analyses of real data in cattle (Woolliams *et al.*, 1994). Embryos were collected every 2 months (selected donors were flushed three times in each evaluation period). The total expected number of transfers per year was 30·6 × N_D. The proportion of transferred embryos resulting in a live calf was 55%. There were no natural matings (all calves were born from embryo transfer).

Figure 1 shows the expected distribution of the number of embryos collected per donor within one evaluation period (three consecutive flushes). Expected frequencies shown in this figure were obtained from a separate stochastic simulation. Embryo production was simulated using the model described above. The number of donors simulated was 100 000, and the number of flushes per donor was three (which corresponds to one evaluation

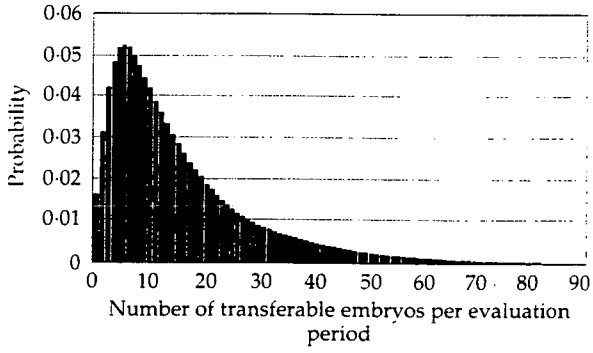


Figure 1 The expected distribution of the number of transferable embryos per donor in one evaluation period (three flushes per donor).

period). Villanueva *et al.* (1995) have presented frequencies for a single flush. The proportion of donors given no embryos was higher than 20% when a single flush was considered.

Comparison among breeding schemes

Schemes were compared at a fixed expected number of calves born per year (N_{CB}). Thus the number of (genetic as opposed to recipient) dams used in conventional schemes was substantially higher than the number of dams used in MOET schemes. The simulation program was run for three different nucleus sizes corresponding to 9, 18 and 36 donors in MOET schemes. The expected number of calves born per year (N_{CB}) were 152, 303 and 606, respectively (in MOET schemes, $N_{CB} = N_D \times 5{\cdot}1 \times 6 \times 0{\cdot}55$). The corresponding numbers of dams used per evaluation period in conventional schemes (*C*) were 147, 286, and 559. The number of selected sires in each evaluation period was six, 12 or 18.

Average true breeding values (G_i) and inbreeding coefficients (F_i) of individuals born at the *i*th year were obtained. Annual rate of response between years *j* and *i* was calculated as $\Delta G_{i-j} = (G_j - G_i)/(j - i)$, where $j > i$. Rates of inbreeding were obtained every year as $\Delta F_i = (F_i - F_{i-1})/(1 - F_{i-1})$ The rate of inbreeding between years *j* and *i* (ΔF_{i-j}) was obtained by taking the average of annual rates. Individual components of annual response (selection intensities and generation intervals for males and females) were also computed and averaged for animals born every year. All breeding schemes were replicated 200 times and results were averaged over all replicates. Schemes were compared for average rates of response and inbreeding from year 15 to year 25 (ΔG_{15-25} and ΔF_{15-25}) and for cumulative response (G_{25}) and inbreeding at year 25 (F_{25}). Average rates of response and inbreeding from year 5 to year 15 (ΔG_{5-15} and ΔF_{5-15}) were also obtained.

Results

In MOET schemes different combinations of selection and mating procedures were simulated: (1) selection on standard BLUP and nested mating design (*M*); (2) selection on standard BLUP and factorial mating design (*MF*); and (3) selection on modified BLUP and factorial mating design (*MFM*). Trends in genetic gain and inbreeding level are shown in Figure 2 for schemes with $N_{CB} = 303$ and N_S

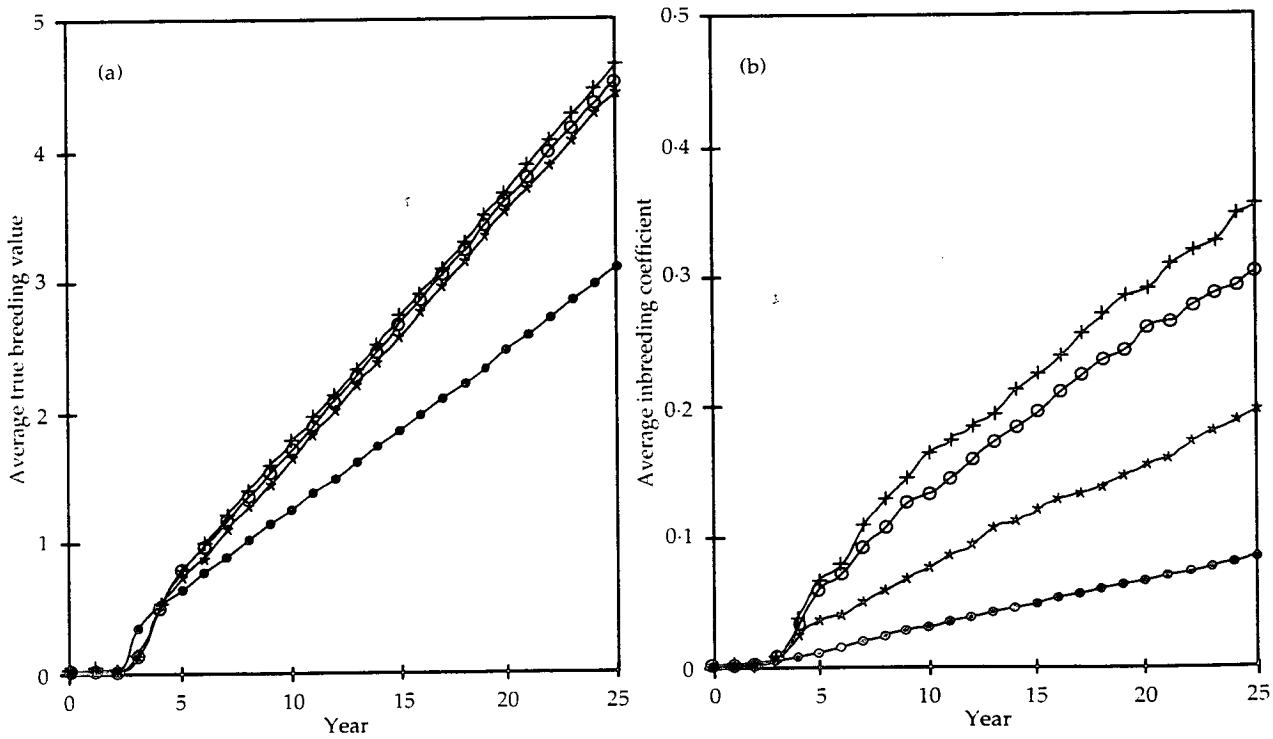


Figure 2 Change in genetic mean (phenotypic standard deviation units) and inbreeding coefficient (%) over years in schemes with $N_{CB} = 303$ and $N_S = 18$: *C* = conventional (◉); *M* = MOET using nested design and standard BLUP (+); *MF* = MOET using factorial design and standard BLUP (○); *MFM* = MOET using factorial design and modified BLUP(✱).

Table 1 *Cumulative genetic progress to year 25 (G_{25}) and average annual rates of genetic progress in early (ΔG_{5-15}) and later years (ΔG_{15-25}) in phenotypic standard deviation units in a closed beef nucleus herd for different reproductive rates and nucleus sizes*†

		N_{CB} = 152	N_{CB} = 303			N_{CB} = 606		
	‡	N_S = 6	N_S = 6	N_S = 12	N_S = 18	N_S = 6	N_S = 12	N_S = 18
G_{25}	*C*	3·12 (0·02)	3·51 (0·02)	3·28 (0·02)	3·06 (0·01)	3·85 (0·02)	3·63 (0·02)	3·46 (0·01)
	MFM	4·21 (0·03)	4·80 (0·03)	4·58 (0·02)	4·37 (0·02)	5·30 (0·03)	5·05 (0·02)	4·89 (0·02)
ΔG_{5-15}	*C*	0·126 (0·001)	0·142 (0·001)	0·132 (0·001)	0·122 (0·001)	0·158 (0·001)	0·146 (0·001)	0·138 (0·001)
	MFM	0·181 (0·002)	0·203 (0·002)	0·192 (0·001)	0·185 (0·001)	0·224 (0·002)	0·211 (0·002)	0·206 (0·001)
ΔG_{15-25}	*C*	0·120 (0·001)	0·139 (0·001)	0·128 (0·001)	0·120 (0·001)	0·150 (0·001)	0·143 (0·001)	0·136 (0·001)
	MFM	0·170 (0·002)	0·196 (0·002)	0·191 (0·002)	0·182 (0·002)	0·220 (0·002)	0·211 (0·001)	0·202 (0·001)

† Standard errors are shown in brackets.
‡ *C* = conventional; *MFM* = MOET using factorial designs and modified BLUP.

= 18. The initial rate of response was reduced substantially due to the Bulmer effect. Annual rates of inbreeding were approximately constant after year 15. Similar trends were observed for different nucleus sizes. The use of MOET (*M*) increased genetic gain by 50% with respect to the conventional scheme. However, this gain was accompanied by a high increase in inbreeding (around 300%). Figure 2 shows that the use of factorial designs and modified BLUP are useful strategies for decreasing inbreeding with a minimal effect on genetic gain.

Cumulative genetic gain and rates of response are shown in Table 1 for conventional (*C*) and MOET schemes using factorial designs and modified BLUP (*MFM*). As expected, rates of response decreased, in general, with years of selection and with number of sires used. Average rates of response from year 15 to year 25 were from 0% to 6% lower than rates of response from year 5 to year 15. By increasing N_S from six to 18 the rate of gain (ΔG_{15-25}) was decreased up to 14% in conventional schemes and by up to 9% in MOET schemes. At a fixed number of sires, by doubling the size of the nucleus the rate of response was increased by between 8% and 16%.

Table 2 shows the rates of inbreeding obtained. In addition, the average inbreeding coefficient at year 25 is shown for comparison. With factorial designs and modified BLUP, the inbreeding rate (ΔF_{15-25}) was 46% to 128% higher in MOET than in conventional schemes. At a fixed number of sires, by approximately doubling the size of the nucleus the rate of inbreeding was decreased by between 15% and 18% in MOET schemes. The effect of the number of sires used on the rate of inbreeding obtained was greater than the effect of the nucleus size. In MOET schemes, the increase in number of sires from six to 18 decreased the rate of inbreeding by nearly 40%.

Thus, the use of MOET leads to increased gains but also to increased inbreeding. When schemes are compared at a constant size (fixed N_{CB}) and at a similar rate of inbreeding, Tables 1 and 2 show that

Table 2 *Average inbreeding coefficient at year 25 (F_{25}) and average annual rates of genetic progress in early (ΔF_{5-15}) and later years (ΔF_{15-25}) in % in a closed beef nucleus herd for different reproductive rates and nucleus sizes*†

		N_{CB} = 152	N_{CB} = 303			N_{CB} = 606		
	‡	N_S = 6	N_S = 6	N_S = 12	N_S = 18	N_S = 6	N_S = 12	N_S = 18
F_{25}	*C*	19·30 (0·26)	17·59 (0·21)	11·76 (0·16)	8·56 (0·11)	17·57 (0·21)	11·42 (0·16)	8·57 (0·12)
	MFM	35·06 (0·47)	29·88 (0·40)	23·80 (0·38)	19·88 (0·27)	25·64 (0·32)	19·72 (0·33)	16·56 (0·28)
ΔF_{5-15}	*C*	0·92 (0·02)	0·81 (0·02)	0·55 (0·01)	0·39 (0·01)	0·81 (0·02)	0·53 (0·01)	0·40 (0·01)
	MFM	1·69 (0·05)	1·44 (0·04)	1·14 (0·03)	0·94 (0·02)	1·22 (0·03)	0·91 (0·03)	0·78 (0·02)
ΔF_{15-25}	*C*	0·90 (0·02)	0·83 (0·02)	0·54 (0·01)	0·39 (0·01)	0·84 (0·01)	0·52 (0·01)	0·39 (0·01)
	MFM	1·76 (0·05)	1·45 (0·04)	1·14 (0·03)	0·89 (0·03)	1·23 (0·03)	0·92 (0·03)	0·74 (0·02)

† Standard errors are shown in brackets.
‡ *C* = conventional; *MFM* = MOET using factorial designs and modified BLUP.

Table 3 *Selection intensities in males (i_{σ}) and females (i_{ς}) in a closed beef nucleus herd for different reproductive rates and nucleus sizes*†

		N_{CB} = 152	N_{CB} = 303			N_{CB} = 606		
	‡	N_S = 6	N_S = 6	N_S = 12	N_S = 18	N_S = 6	N_S = 12	N_S = 18
i_{σ}	*C*	1·53 (0·004)	1·85 (0·004)	1·53 (0·003)	1·33 (0·002)	2·16 (0·004)	1·86 (0·003)	1·67 (0·003)
	MFM	1·51 (0·006)	1·86 (0·005)	1·53 (0·003)	1·33 (0·003)	2·18 (0·005)	1·86 (0·003)	1·67 (0·003)
i_{ς}	*C*	0·29 (0·002)	0·28 (0·001)	0·29 (0·001)	0·30 (0·001)	0·27 (0·001)	0·28 (0·001)	0·29 (0·001)
	MFM	1·39 (0·005)	1·38 (0·003)	1·39 (0·003)	1·40 (0·003)	1·38 (0·003)	1·38 (0·003)	1·39 (0·002)

† Standard errors are shown in brackets.
‡ *C* = conventional; *MFM* = MOET using factorial designs and modified BLUP.

the extra gain from MOET is around 30% (for instance, for N_{CB} = 303, when comparing ΔG_{15-25} for *MFM* with N_S = 18 to *C* with N_S = 6, the relative advantage of MOET is 31%).

Male and female selection intensities and generation intervals are shown in Tables 3 and 4, respectively. Values presented are averages over the last 15 years of selection. The extra gains obtained with MOET schemes were mostly a consequence of the increased female selection intensity. This increased by 367% to 411% with respect to conventional schemes. Male selection intensity was very similar with or without MOET. The use of MOET also led to reductions in female generations intervals (Table 4). These were reduced by around 11% with respect to conventional schemes. Male generation intervals were increased (between 0% and 7%) in MOET schemes, due to the selection criterion used (*M* and *MF* schemes produced the same male generation intervals, which were lower than that with *C*). With modified BLUP, by increasing the heritability, the weight attached to the Mendelian sampling term of the individual is increased with respect to standard BLUP. Mendelian sampling terms of individuals with a large amount of information on their own or their offspring performance are estimated more accurately. Then, older animals (animals with a large amount of offspring) will be favoured when using modified BLUP.

Discussion

Alternative MOET breeding schemes for beef cattle were examined in this study, with the aim of finding those which may be practicable. The results presented here clearly demonstrate the effects of nucleus size and sire selection intensity on expected genetic gain in conventional and MOET schemes. Expected genetic gain in MOET schemes was substantially higher than that in conventional schemes of equivalent size, but was accompanied by higher levels of inbreeding. However, as expected (Woolliams, 1989; Grundy *et al.*, 1994), these increases in inbreeding were limited by the use of factorial mating, especially in conjunction with modified BLUP selection. As reported in another study (Villanueva *et al.*, 1994) although modified BLUP selection led to a slight reduction in expected genetic gain, compared to conventional BLUP selection, this was more than compensated for by a substantial reduction in inbreeding.

Many methods have been proposed to decrease rates

Table 4 *Generation intervals for males (L_{σ}) and females (L_{ς}) in years in a closed nucleus herd for different reproductive rates and nucleus sizes*†

		N_{CB} = 152	N_{CB} = 303			N_{CB} = 606		
	‡	N_S = 6	N_S = 6	N_S = 12	N_S = 18	N_S = 6	N_S = 12	N_S = 18
L_{σ}	*C*	3·31 (0·015)	3·01 (0·010)	3·27 (0·010)	3·48 (0·011)	2·85 (0·008)	3·01 (0·007)	3·16 (0·007)
	MFM	3·48 (0·018)	3·22 (0·013)	3·36 (0·011)	3·46 (0·011)	3·03 (0·010)	3·15 (0·008)	3·22 (0·015)
L_{ς}	*C*	4·11 (0·011)	3·89 (0·007)	4·04 (0·006)	4·19 (0·008)	3·74 (0·005)	3·86 (0·005)	3·96 (0·005)
	MFM	3·66 (0·019)	3·51 (0·010)	3·55 (0·010)	3·61 (0·011)	3·38 (0·007)	3·44 (0·007)	3·47 (0·014)

† Standard errors are shown in brackets.
‡ *C* = conventional; *MFM* = MOET using factorial designs and modified BLUP.

of inbreeding in breeding programmes (see e.g. Toro and Perez-Enciso, 1990). In general, their use will also lead to decreases in gain. The strategies used in this paper (factorial mating designs and modified BLUP) are simple to use and very efficient in the sense that they lead to considerable reductions in inbreeding with very little loss in response. Other methods reduce inbreeding with minimal effect on gain (Brisbane and Gibson, 1994; Wray and Goddard, 1994) but their use is more complicated. Results for different nucleus sizes are only presented for MOET schemes using strategies for controlling inbreeding (Tables 1 and 2). However, simulations were also run for MOET schemes using different combinations of selection criteria and mating designs (*M* and *MF*). For different nucleus sizes, the use of strategies for controlling inbreeding in MOET schemes resulted in a decrease of 32% to 46% in the rate of inbreeding. Inbreeding was still high in these MOET schemes but the measures for reducing inbreeding were clearly effective. This is also illustrated in Figure 2. The smaller the number of sires the smaller the benefit of using strategies for decreasing inbreeding. Most of the reduction in inbreeding was due to the modified selection criterion. The decrease in inbreeding rates with factorial designs ranged from 1% to 27%. When modified BLUP was the selection criterion, generation intervals were increased by 10% to 17%. This led to small decreases in genetic gain in *MFM* schemes.

Wray and Simm (1990) reported lower rates of inbreeding for equivalent schemes to those presented here. For example, for heifer embryo transfer (ET) schemes (heifers flushed once at 15 months of age) and N_{CB} = 200 and 400 the annual inbreeding rates under standard BLUP selection were 2·11 and 1·65%, respectively, when five sires were used. Rates of inbreeding were decreased in both schemes to 0·95 when the number of sires was increased to 15. Rates of inbreeding were lower in adult schemes under standard BLUP and factorial mating designs. In our study, for N_{CB} = 303 the rate of inbreeding in schemes *M* was 2·31, 2·22 and 1·82% for six, 12 and 18 sires, respectively. With factorial designs (*MF*), inbreeding rates were 2·28, 1·62 and 1·44% for six, 12 and 18 sires, respectively. The coefficient of variation and the repeatability of embryo yield were shown to have an important effect on the inbreeding obtained (Villanueva *et al.*, 1995). Previous studies evaluating MOET schemes have assumed zero correlation between the numbers of embryos obtained in successive flushes and lower coefficient of variation than that used in this study. This can explain (at least in part) the higher rates of inbreeding obtained for MOET schemes here. Since commercial animals are often crossbreds, the problem of inbreeding depression in beef cattle can be limited (Gama and Smith, 1993). However, inbreeding can limit the genetic improvement carried out in nucleus and other pedigree herds.

Deciding on the most appropriate basis to compare conventional and MOET schemes is difficult. The problem of additional inbreeding in MOET schemes could be minimized by considerably increasing the size of the population. However, this would increase the costs proportionally and a compromise is needed. Deciding on the advisable maximum level of inbreeding in practical situations is also very difficult. If, for the purpose of discussion, we assume that a rate of 1% per year is acceptable then MOET schemes would be genetically appropriate providing their size is large enough. The smallest MOET scheme considered (N_{CB} = 152) gave an unacceptable level of inbreeding even when modified BLUP and a factorial design were used (Table 2). In the schemes presented, animals were allowed to be selected for several evaluation periods. Villanueva *et al.* (1994) have showed that culling of females from the herd after having been used for one evaluation period is an efficient strategy for decreasing inbreeding (although response is also affected). When the use of selected males is also limited to one period, inbreeding is further reduced with no further decrease in response. Culling strategies were applied to MOET schemes with N_{CB} = 152 and results indicate that limiting the use of both males and females leads to acceptable rates of inbreeding (around 1% per annum). In this case, the advantage over the conventional scheme was around 30% (results not shown).

MOET schemes with 18 donors using 18 sires, modified BLUP and factorial designs gave an annual rate of inbreeding lower than 1%. Compared with a conventional scheme of the same size, this scheme gave about 50% more genetic gain but also a substantially higher rate of inbreeding (about 130%). MOET schemes with 36 donors would lead to rates of inbreeding up to 90% higher than those in corresponding conventional schemes. It could be argued that the fairest basis for comparison of schemes is at a similar size and at a similar acceptable rate of inbreeding. On this basis, the results in Tables 1 and 2 show that MOET schemes employing effective methods of controlling inbreeding could achieve about 30% higher genetic gain than conventional schemes of equivalent size. This advantage of MOET is lower than that previously reported. Wray and Simm (1990) found an advantage of MOET over conventional schemes of around 50% in beef cattle when schemes were compared at the same level of inbreeding. Bondoc and Smith (1993) carried out equivalent comparisons

between adult MOET and conventional progeny testing schemes in dairy cattle. The benefit from MOET was around 50% for a trait with heritability 0·25.

In this study, schemes have been compared at a fixed expected number of calves born per year. With current parameters for embryo production, limiting the number of embryos transferred from a given selected dam to reduce inbreeding will decrease the total number of embryos transferred (and the total number of calves born). The total number of calves born could be maintained by increasing the number of flushes in cows with low embryo yield. However, this would increase generation intervals (and consequently it would decrease response). In order to avoid an increase in generation intervals, improved embryo technologies (increased embryo yields) are needed. Villanueva *et al.* (1995) have found that with improved techniques inbreeding rates can be substantially reduced due to the opportunity of equalizing family sizes.

Nicholas (1989) examined the risk of the scheme in terms of variability of response to selection. He found that restricting the coefficient of variation of response to acceptable levels (5 to 10% after 10 years of selection) gives acceptable levels of inbreeding. In our study, only the larger schemes (N_{CB} = 606) with $N_S \geq 12$ gave coefficients of variation of gain smaller than 10%.

MOET nucleus schemes have often been proposed as alternatives to breeding programmes depending on national recording schemes. However, the conventional schemes simulated here are essentially nucleus schemes with natural mating. It is very difficult to model breeding programmes in the wider population for beef cattle because, at least in many European countries, the level of recording is very low and across herd evaluations have only recently become available. However, comparisons between MOET schemes and larger conventional schemes (as conducted in this study) give an indication of the opportunity of increasing response by selecting from a larger population. These comparisons illustrate that there is still a benefit from MOET schemes. For example, MOET schemes with N_{CB} = 303 and N_S = 18 give similar inbreeding rates to conventional schemes with N_{CB} = 606 and N_S = 6 (Table 2). The benefit from MOET with respect to rate of gain is 21% in this case (Table 1). However, comparisons would be unfair if infinite resources are assumed in conventional schemes. In the extreme, comparisons would be made between finite and infinite populations. Finally, there can be benefits from maintaining an open nucleus when considering not only gain but also inbreeding. In dairy cattle MOET schemes, opening the nucleus to importation of progeny tested males has been shown to be an effective method of controlling inbreeding with small effect on gain (e.g. Meuwissen, 1991). So in practice, both MOET and conventional schemes would benefit from access to a wider recorded population.

Acknowledgements

This work was funded by the Ministry of Agriculture, Fisheries and Food, the Milk Marketing Board of England and Wales and the Meat and Livestock Commission. The Scottish Agricultural College receives financial support from the Scottish Office Agriculture and Fisheries Department. We would like to thank Dr B. J. McGuirk for his comments on the manuscript.

References

Bondoc, O. L. and Smith, C. 1993. Optimized testing schemes using nucleus progeny, adult MOET siblings, or juvenile MOET pedigrees in dairy cattle closed populations. *Journal of Animal Breeding and Genetics* **110:** 30-40.

Bourdon, R. M. and Brinks, J. S. 1982. Genetic, environmental and phenotypic relationships among gestation length, birth weight, growth traits and age at first calving in beef cattle. *Journal of Animal Science* **55:** 543-553.

Brisbane, J. R and Gibson, J. P. 1994. Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 19,* pp. 135-138.

Broadbent, P. J. 1990. The NOSCA Simmental MOET project. *British Cattle Breeders Club, Winter Conference, Cambridge.*

Gama, L. T. and Smith, C. 1993. The role of inbreeding depression in livestock production systems. *Livestock Production Science* **36:** 203-211.

Grundy, B., Caballero, A., Santiago, E. and Hill, W. G. 1994. Use of biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. *Animal Production* **59:** 465-468.

Land, R. B. and Hill, W. G. 1975. The possible use of superovulation and embryo transfer in cattle to increase response to selection. *Animal Production* **21:** 1-12.

Luo, Z. W., Woolliams, J. A. and Simm, G. 1994. An assessment of present and future effectiveness of embryological techniques in ruminants. *AgBiotech News and Information* **6:** 13N-18N.

Meuwissen, T. H. E. 1991. Expectation and variance of genetic gain in open and closed nucleus and progeny testing schemes. *Animal Production* **53:** 133-141.

Nicholas, F. W. 1979. The genetical implications of multiple ovulation and embryo transfer in small dairy herds. *Proceedings of the thirtieth annual meeting of the European Association of Animal Production, Harrogate.*

Nicholas, F. W. 1989. Incorporation of new reproductive technology in genetic improvement programmes. In *Evolution and animal breeding* (ed. W. G. Hill and F. C. MacKay), pp. 203-209. CBA International, Wallingford.

Nicholas, F. W. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

Quaas, R. L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* **32:** 949-953.

Scott, D. 1992. Accelerated development of a British beef breed. *British Cattle Breeders Club, Winter Conference, Cambridge.*

Simm, G., Steane, D. E. and Wray, N. R. 1990. Developments in beef cattle breeding programmes in Europe. *Proceedings of the fourth world congress on genetics applied to livestock production, Edinburgh, vol. XV,* pp 231-243.

Toro, M. and Perez-Enciso, M. 1990. Optimization of selection response under restricted inbreeding. *Genetics Selection Evolution* **22:** 93-107.

Villanueva, B. and Simm, G. 1994. The use and value of embryo manipulation techniques in animal breeding. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 20,* pp. 200-207.

Villanueva, B., Woolliams, J. A. and Simm, G. 1994. Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics Selection Evolution* **26:** 517-535.

Villanueva, B., Woolliams, J. A. and Simm, G. 1995. The effect of improved reproductive performance on genetic gain and inbreeding in MOET breeding schemes for beef cattle. *Genetics Selection Evolution* **27:** 275-285.

Woolliams, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49:** 1-14.

Woolliams, J. A., Luo, Z. W., Villanueva, B., Waddington, D., Broadbent, P., McKelvey, W. A. C. and Robinson, J. J. 1994. Analysis of factors affecting superovulatory responses in ruminants. *Journal of Agricultural Science, Cambridge* **124:** 61-70.

Wray, N. R. and Goddard, M. E. 1994. Increasing long-term response to selection. *Genetics Selection Evolution* **26:** 431-451.

Wray, N. R. and Simm, G. 1990. The use of embryo transfer to accelerate genetic improvement in beef cattle. *Proceedings of the fourth world congress on genetics applied to livestock production, Edinburgh, vol. XV,* pp. 315-318.

(*Received* 14 *November* 1994—*Accepted* 3 *April* 1995)

Paper 23

Animal Science 1995, 60: 379-387

0003-3561/95/43910379$20-00

Controlling inbreeding in dairy MOET nucleus schemes

Z. W. Luo, J. A. Woolliams† and R. Thompson

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS

Abstract

A nucleus dairy population using multiple ovulation and embryo transfer (MOET) was stochastically modelled with overlapping generations. The aim was to investigate the feasibility of controlling inbreeding in MOET breeding schemes using more realistic parameters for embryo recovery and best linear unbiased prediction (BLUP) for genetic evaluation. Four different cases (involving the culling of donors, more donors and the use of organized progeny testing of nucleus bulls) were studied in combination with nested and factorial designs. Further studies involved modifications of the selection index, including subtracting parental breeding values, inflating the genetic variance in the BLUP evaluation and penalizing inbred animals; these options were examined both with and without organized progeny testing. The effects of applying these schemes on both genetic response and rate of inbreeding were investigated. The results stressed the importance of incorporating progeny testing into MOET schemes for value of reducing inbreeding whilst maintaining genetic progress. There was no significant difference between nested and factorial designs. In the absence of progeny testing the inflation of genetic variance was more effective than subtracting parental breeding values at controlling inbreeding; however incorporating progeny testing made the latter strategy more potent and the superiority of inflating the genetic variance was in this case much smaller and non-significant.

Keywords: *dairy cattle, inbreeding, MOET.*

Introduction

Incorporation of multiple ovulation and embryo transfer (MOET) into selection programmes for dairy cattle was previewed by the results of Nicholas and Smith (1983) and Colleau (1985). Subsequently expectations of the benefits of closed nucleus schemes using MOET were reduced following simulation studies (e.g. Juga and Maki-Tanila, 1987; Ruane and Thompson, 1991). However with more sophisticated models Meuwissen (1991) suggested increased reproductive rates in females would increase genetic gain 1·08- to 1·16-fold. Compared with conventional progeny testing the genetic gain was achieved through higher selection differentials in females and shorter generation intervals, the latter usually arising from the sacrifice of selection accuracy, particularly in bulls. Colleau (1985) also emphasized that the advantage in genetic progress of using MOET depended on the size of nucleus and on the technical parameters of transfer such as embryo number and survival rate.

† The author for correspondence.

These studies above have mainly concentrated on investigating the potential advantages of using the modern reproductive techniques to obtain extra genetic improvement. However, the higher level of inbreeding in the breeding schemes that are usually encountered has had less emphasis. Inbreeding can directly cause a decrease in genetic variation, hence decrease in response to selection and, through inbreeding depression, a decrease in performance traits and population fitness. Therefore, it is important to assess the level of inbreeding likely to be encountered in MOET schemes and to take steps to control it.

Woolliams (1989) analysed rates of inbreeding for modified MOET schemes with factorial or hierarchical mating, and found that the use of the factorial mating designs in the breeding schemes could reduce rate of inbreeding, whilst avoiding any loss in genetic response. Strandén, Maki-Tanila and Mäntysaari (1991) included the effects of mating strategies in their study of inbreeding rates in MOET schemes and found that factorial schemes reduced

inbreeding, or if not were yielding higher rates of genetic progress. Subsequently de Boer and van Arendonk (1994) and Leitch, Smith, Burnside and Quinton (1995) have shown clear advantages of factorial designs when comparing rates of genetic progress at the same rate of inbreeding.

Bondoc and Smith (1993) looked at the potential of progeny testing to reduce inbreeding in the nucleus and compared inbreeding rates with progeny testing to those of adult and juvenile MOET schemes with each producing their maximal rate of gain. Their results indicated that progeny testing proportionately reduced the annual inbreeding rate to only 0·1 of that for the other schemes. However, in these models progeny testing was restricted to being within a nucleus of fixed test capacity and so the progeny testing scheme also produced low rates of genetic progress compared with more standard models of progeny testing.

Other studies that have used best linear unbiased predictors (BLUP) for evaluation have investigated the possibilities of defining selection criteria based upon the components of these predictions in order to reduce inbreeding. In this context Verrier, Colleau and Foulley (1993) examined the potential of subtracting parental information from the estimated breeding value, and other alternatives have been suggested such as biasing the heritability upwards in the evaluation (Grundy and Hill, 1993).

A problem with many of these studies is that they include less than realistic assumptions concerning the breeding scheme and the biology of MOET. In particular the operation of a scheme with evaluation using BLUP is likely to be run with overlapping generations in both sexes arising from selection across cohorts. Further the studies have often assumed an inappropriate distribution of embryo yields: evidence (Lohuis, Smith and Dekkers, 1993; Woolliams, Luo, Villanueva, Waddington, Broadbent, McKelvey and Robinson , 1994) suggests that for a mean number of transferable embryos per recovery close to five, the coefficient of variation between recoveries is of the order of 1·2 (cf. a poisson with a mean of five for which the CV is 0·45) and the number has a repeatability of 0·2 across donor cows. Both these characteristics will act to increase the variance of family size, and hence inbreeding, in a breeding scheme.

The present study had as its objective the investigation of strategies for controlling inbreeding and maintaining genetic gain in MOET nucleus schemes for dairy cattle, using stochastic simulation studies that incorporated reproductive parameters obtained from a recent extensive survey (Luo, Woolliams and Sim, 1994). These strategies concern the management of donors, the progeny testing of bulls, mating design and the use of information from BLUP.

Methods

Genetic model

The objective trait (e.g. milk or milk solids yield) was assumed to have an inheritance described by the additive infinitesimal model, with a heritability of 0·4, a repeatability of 0·6 and a phenotypic variance of 1·0 in an unselected population. Only females had performance information, but this might include repeated lactation records and part-lactations.

In the initial generation the founders were assumed unselected and unrelated and their true breeding values (TBV) were generated from a normal distribution with mean 0 and variance 0·4. Subsequently the true breeding value of a newborn animal was generated as the average of the TBV of the parents plus a normally distributed Mendelian sampling term with variance $(1-\bar{F})\sigma_{a,0}^2/2$, where $\bar{F}$ is the average of the inbreeding coefficients of the parents and $\sigma_{a,0}^2$ is the initial additive genetic variance. The repeatable component of the phenotypes was obtained for each cow by adding a normal random deviate of mean zero and variance 0·2 to the TBV and the phenotype itself was obtained for each lactation by further adding a normal deviate of mean zero and variance 0·4.

The estimated breeding value (EBV) of an animal was obtained using BLUP from an individual animal model. The model fitted was

$$y_{ij} = \mu + \alpha_i + \gamma_i + \varepsilon_{ij}$$

where y_{ij} represents record j of animal i, and μ, α_i, γ_i, and ε_{ij} respectively represent the overall mean, the breeding value and permanent environmental effect of animal i and residual deviation respectively; μ was considered as a fixed effect whilst the remainder were treated as random. Account was taken in the BLUP evaluation of animals whose full lactation record was not completed when their EBV evaluation was updated, using the method of Wiggans, Misztal and Van Vleck (1988).

Basic breeding scheme (case 1)

Structure and operation. The founders of the nucleus herd were males and females in three age groups (2 to 4 years old). The herd was operated with overlapping of generations and no specific assumptions were made about the age distribution of the nucleus population at any time. Thus the female population simply consisted of three groups:

immature females, mature non-pregnant females and pregnant females. All information on breeding values for both sexes was assumed to come from the observations made within the nucleus.

Selection was carried out for 25 years, which was divided into periods and after each period breeding values were evaluated. In the present study, the evaluation period was 6 months. Time was counted in units of days and animal status (including age and candidacy for breeding, but not estimated breeding value) was updated every day. Animals of all ages were subject to a mortality with survival rates that varied from 0·9950 to 0·9500 per period and decreased with age.

Males and females were selected from eligible candidates (mature but not pregnant) based upon their EBVs irrespective of whether they have been selected in a previous period. The top nine bulls were selected for use as sires in all services. The 18 cows with the highest EBV were used as embryo transfer (ET) donors and a further 126 cows were selected and received up to three artificial inseminator (AI) services with an overall success rate of 98%. These cows as well as producing additional offspring also served to provide lactation records and hence information for the genetic evaluation of the nucleus animals. ET recipients were assumed to be provided from outside the nucleus. Those candidate cows that remained unselected and those that were selected but failed to conceive to AI were culled. The number of ET donors selected here mimics the size of the ET nucleus herd established in the UK and which carried out an average of 1000 transfers per year; while the number of bulls was chosen for convenience with regard to the mating designs and the intended number of recoveries. Semen was assumed to be available from all previously selected bulls. The selected donors were mated to selected sires according to a specified design (see later).

Reproductive performance. Males were assumed capable of breeding at 12 months of age and females were assumed to be ready for breeding and superovulation at 15 months of age. Recoveries were assumed to be possible three times in a 6-month period. The number of transferable embryos collected from a single recovery from an individual cow was derived from a negative binomial distribution. A repeatability of 0·23 between cows over recoveries was modelled following the method of Woolliams *et al.* (1994). The mean and coefficient of variation of numbers of embryos obtained in a single recovery from a nucleus cow was 5·0 and 1·3 respectively. These parameters are very close to those from analyses of extensive data on embryo recovery (Woolliams *et al.*, 1994). Embryos survived until calving with a fixed probability of 0·55 in accord with survey results (Luo *et al.*, 1994).

Alternative selection strategies

The rate of inbreeding is proportional to the expected sum of squared long-term contributions of an ancestor to its descendants (Wray and Thompson, 1990). Theoretical analysis (Woolliams, Wray and Thompson, 1993) has shown that the inbreeding rate can be partitioned into three components defined in terms of the long-term contributions: (i) the squared mean; (ii) their sampling variance ignoring full-sibs; and (iii) the sampling covariance from full-sibs. Their sum in a population under selection contains as one of several terms the rate of inbreeding estimated from the variance of family size in one generation (Hill, 1979). Strategies to reduce variance in one generation family sizes might therefore reduce inbreeding in selection. For a given number of transfers, the variance is the least when selected animals in each of the two sexes contribute equally and it increases as fewer females are selected for use as donors or fewer males are selected. The standard scheme described above was compared with three modified schemes (cases 2 to 4; not mutually exclusive) each with two alternative mating designs, which will have differing effects on one or more of these components.

Case 2: culling used donors. One method of restricting the variance of family size is to restrict the opportunity of contributing. Therefore donors were culled after three recoveries. This is equivalent to a restriction upon the expected maximum physical family size of a donor.

Case 3: using more donors. The expected physical family size of ET donors can also be controlled by selecting more females to supply the same number of embryos in each period. In the basic scheme (described above) each of the 18 donors had three recoveries of embryos and this was compared with selecting 27 donors in each evaluation period with each donor having only two recoveries. The expected total number of embryos collected in any one period remained the same.

Case 4: including progeny testing. Progeny testing outside the nucleus herd introduces the potential for accurate estimates of breeding values for males. Further, since the method relies principally upon offspring rather than pedigree, the correlations of estimated breeding values among full- and half-sibs will be lower than if derived solely from nucleus information. This reduces the total sampling variance of the long-term contributions. Progeny testing to a given precision (as measured by the number of

daughter lactations) was incorporated into the simulations using the reduced animal model of Quaas and Pollak (1980). In the simulated progeny-testing schemes, all bulls of 6 years old were assumed to be progeny tested with 50 daughters. Once a bull was progeny tested, its estimated breeding value was held constant in further evaluations.

Mating designs. Two mating designs (factorial and hierarchical) were investigated. In each period it was impossible to obtain a strictly balanced factorial design due to the restricted number of recoveries. Two rules were thus followed in the present study in order to obtain a factorial design: (i) every selected bull was used an equal number of times; and, (ii) no bull was mated to the same donor more than once in a period. This was able to be satisfied for the numbers of sires and ET donors considered in the present study. After following these rules bulls were randomly assigned to donors. The factorial design was intended to reduce the component of inbreeding due to the sampling covariance amongst full-sibs.

Selecting upon modified indices
In the schemes described above, selection of a candidate was upon its EBV. Modifications of this index were considered which were functions of the EBVs of the population and these were examined using factorial designs with the basic scheme with and without progeny testing (cases 1 and 4).

Partial subtraction of parental breeding values (I_1). Index (I_1) was derived for an individual by partial subtraction of the estimated breeding value of the sire and dam from its own estimated breeding value (Verrier *et al.*, 1993), i.e. $I_{1,i} = EBV_i - \lambda_s EBV_s - \lambda_d EBV_d$. When $\lambda_s = \lambda_d = 1/2$ the scheme is selecting only upon the estimated Mendelian sampling component of the breeding value, which is a form of within-family selection, while with $\lambda_s = \lambda_d = 0$, the index is not modified and is equal to the best linear unbiased predictor.

Inflating the genetic variance in BLUP evaluation (I_2). The lower the heritability of the quantitative trait, the more important the family information is in the evaluation and thus a higher rate of inbreeding is expected. Grundy and Hill (1993) examined the use of a heritability which is artifically biased upwards in the evaluation so as to restrict family information in BLUP selection. Index I_2 was derived by artificially inflating proportionately the genetic variance by 0·25 or 0·75 in the BLUP evaluations. It is an open question whether fixed effects should be estimated and data corrected using the actual or inflated genetic variance, but in this study with only one fixed effect little difference was expected.

Penalizing inbred individuals (I_3). A further index $I_3 = (1 - F_i)\, EBV_i$ was used, where F_i represents the inbreeding coefficient of candidate i. This latter form was motivated by the observation that a relatively high inbreeding coefficient of an individual indicated a more restricted range among its own ancestors. Whilst the range of ancestors in the population is a property of the selected group as a whole, this might prove a relatively simple way of encouraging a spread of contributors.

Method for comparison of the breeding schemes
Each combination of the four cases and the two mating designs were replicated 20 times. In the analysis of these results there was a pooling across mating designs and cases, thus results presented represent 40 replicates for each case and 80 replicates for each mating design. All modified indices of the form I_1 and I_2 were replicated 20 times both with and without progeny testing, and a subset of parameters were replicated 40 times. I_3 was replicated 40 times with and without progeny testing. Running the simulation of these breeding schemes was computationally very demanding, e.g. it took an average of more than 6 h in CPU to run a single scheme with this number of replicates on a VAX 4600 main frame computer.

Annual genetic response at year t for each replicate, in phenotypic standard deviation units, was calculated as $\Delta G_t = TBV_t - TBV_{t-1}$ where TBV_t represents the mean of true breeding values of animals born in year t. The mean annual genetic response from year 15 to year 25 (ΔG) was calculated by averaging ΔG_t ($16 \leq t \leq 25$) for each replicate.

Annual inbreeding rate at year t for each replicate was calculated as $\Delta F_t = (F_t - F_{t-1})(1 - F_{t-1})^{-1}$, where F_t represents the mean inbreeding coefficient of new born animals at year t. The mean annual inbreeding rate from year 15 to year 25 (ΔF) was calculated for each replicate by averaging ΔF_t for $16 \leq t \leq 25$.

The generation interval for selected males (or females) was calculated as the average age of the parents when the selected individuals were born. Average generation intervals were calculated for both sexes over years 5 to 25.

Other results calculated but not presented were the selection intensity for each sex in each year, the accuracy of evaluation (calculated as the empirical correlation coefficient between true breeding values and estimated breeding values of candidate animals) for each sex, and the additive genetic variance in each year calculated from the true breeding values of animals born in that year.

Table 1 *Rates of genetic progress (ΔG) and inbreeding (ΔF) between years 15 and 25 and generation intervals (L) for options involving selection upon BLUP without modification*

Scheme	ΔG	ΔF	Generation intervals† L_m	L_f	L
Basic scheme	0·1252 (0·0045)	0·0269 (0·0016)	3·00	3·32	3·16
Culling donors	0·1128 (0·0047)	0·0207 (0·0017)	3·57	2·67	3·12
Using more donors	0·1195 (0·0042)	0·0223 (0·0018)	2·88	3·12	3·00
With progeny testing	0·1295 (0·0029)	0·0121 (0·0020)	5·54	2·60	4·07

† s.e.s for L_m and L_f are all less than 0·02 apart from progeny testing where the s.e. of L_m was 0·08.

Results

The results are given in two sections: (i) the options where evaluation and selection is based solely upon BLUP; and (ii) options in which selection is based upon modifying the selection index.

Options involving BLUP evaluation only

Table 1 shows the rates of genetic progress and inbreeding between years 15 and 25 and generation intervals for the basic scheme and the schemes involving progeny testing or closer control of donor family sizes. It can be seen that the most dramatic effect came from the incorporation of progeny testing reducing ΔF to 0·45 of that found in the basic scheme ($P < 0·001$); whilst ΔG increased to 1·03 of that in the basic scheme. Although the increase in ΔG was not significant a symmetric 95% confidence interval for the different in ΔG with and without progeny testing is $(-0·0077\sigma_p, 0·0163\sigma_p)$ i.e. pessimistically a 0·06 proportional decrease, and optimistically a 0·13 proportional increase in ΔG from incorporating progeny testing.

Table 2 *Rates of genetic progress (ΔG) and inbreeding (ΔF), and generation intervals for index I (partial subtraction of parental breeding values) without progeny testing (s.e.s are in parentheses)*

	ΔG	ΔF	Generation intervals† L_m	L_f	L
($\lambda_s = 0, \lambda_d = 0$)	0·1252 (0·0045)	0·0269 (0·0016)	3·00	3·32	3·16
($\lambda_s = 1/4, \lambda_d = 1/4$)	0·1145 (0·0038)	0·0227 (0·0025)	3·82	3·85	3·84
($\lambda_s = 1/8, \lambda_d = 0$)	0·1249 (0·0044)	0·0269 (0·0025)	3·07	3·40	3·24
($\lambda_s = 1/4, \lambda_d = 0$)	0·1204 (0·0060)	0·0212 (0·0023)	3·27	3·56	3·42
($\lambda_s = 1/2, \lambda_d = 0$)	0·1008 (0·0045)	0·0193 (0·0018)	4·73	4·30	4·52
($\lambda_s = 0, \lambda_d = 1/8$)	0·1286 (0·0055)	0·0264 (0·0028)	3·13	3·37	3·25
($\lambda_s = 0, \lambda_d = 1/4$)	0·1147 (0·0050)	0·0229 (0·0023)	3·45	3·51	3·48
($\lambda_s = 0, \lambda_d = 1/2$)	0·0826 (0·0041)	0·0182 (0·0019)	5·17	4·42	4·80

† s.e.s are 0·03 for L_f and ≤0·1 for L_m.

Table 3 *Rates of genetic progress (ΔG) and inbreeding (ΔF), and generation intervals for index I (partial subtraction of parental breeding values) with progeny testing (s.e.s are in parentheses)*

	ΔG	ΔF	Generation intervals† L_m	L_f	L
($\lambda_s = 0, \lambda_d = 0$)	0·1295 (0·0029)	0·0121 (0·0020)	5·54	2·60	4·07
($\lambda_s = 1/4, \lambda_d = 1/4$)	0·1216 (0·0019)	0·0062 (0·0010)	5·08	3·51	4·30
($\lambda_s = 1/8, \lambda_d = 0$)	0·1314 (0·0035)	0·0101 (0·0022)	5·83	3·29	4·56
($\lambda_s = 1/4, \lambda_d = 0$)	0·1342 (0·0046)	0·0086 (0·0022)	6·23	3·51	4·87
($\lambda_s = 1/2, \lambda_d = 0$)	0·1178 (0·0025)	0·0042 (0·0007)	8·02	4·22	6·12
($\lambda_s = 0, \lambda_d = 1/8$)	0·1334 (0·0048)	0·0124 (0·0013)	5·03	3·27	4·15
($\lambda_s = 0, \lambda_d = 1/4$)	0·1304 (0·0030)	0·0087 (0·0019)	6·57	3·46	5·02
($\lambda_s = 0, \lambda_d = 1/2$)	0·1181 (0·0041)	0·0055 (0·0005)	7·79	4·13	5·96

† s.e.s are 0·03 for L_f and ≤0·14 for L_m.

Culling donors, reduced ΔF ($P < 0·05$) to 0·77 of those for the basic scheme and a similar trend was evident for increasing numbers of donors ($P < 0·01$) but the reduction was smaller (0·83). However, culling donors also reduced ΔG ($P < 0·05$) to 0·90 of that in the basic scheme and again increasing the number of donors showed a similar trend only less extreme and non-significant. The results from using factorial or hierarchical mating designs were not significant although factorial designs reduced ΔF by 0·0024 (s.e. 0·0022) and increased ΔG by 0·0014 (s.e. 0·0060).

Modifying the index of selection

Index I_1. Table 2 shows the effect of subtracting pedigree information without progeny testing ratios of ΔG with the basic scheme were 0·96, 0·92 and 0·91 for ($\lambda_s = 1/4, \lambda_d = 0$), ($\lambda_s = 0, \lambda_d = 1/4$) and $\lambda_s = 1/4, \lambda_d = 1/4$) respectively, whilst reductions in ΔF were 0·79, 0·85 and 0·84 respectively. Generation intervals were

Table 4 *Rates of genetic progress* (ΔG) *and inbreeding* (ΔF), *and generation intervals from varying the genetic variation assumed in BLUP evaluation (s.e.s are in parentheses)*

Scheme	ΔG	ΔF	Generation intervals† L_m	L_f	L
Without progeny testing					
$v_a^2 = \sigma_{a,0}^2$	0·1252 (0·0045)	0·0269 (0·0016)	3·00	3·32	3·16
$v_a^2 = \sigma_{a,0}^2 \times 1{\cdot}25$	0·1234 (0·0033)	0·0241 (0·0016)	3·10	3·35	3·23
$v_a^2 = \sigma_{a,0}^2 \times 1{\cdot}75$	0·1359 (0·0051)	0·0190 (0·0014)	3·32	3·37	3·35
With progeny testing					
$v_a^2 = \sigma_{a,0}^2$	0·1295 (0·0029)	0·0121 (0·0020)	5·45	2·60	4·07
$v_a^2 = \sigma_{a,0}^2 \times 1{\cdot}25$	0·1408 (0·0026)	0·0113 (0·0024)	4·16	3·20	3·68
$v_a^2 = \sigma_{a,0}^2 \times 1{\cdot}75$	0·1413 (0·0031)	0·0086 (0·0021)	4·12	3·23	3·68

† s.e.s are ≤0·02 for both L_f and L_m.

increased in both sexes. With progeny testing (see Table 3) the ratios of ΔG to the basic scheme with progeny testing were 1·04 and 1·01 for ($\lambda_s = 1/4$, $\lambda_d = 0$) and ($\lambda_s = 0$, $\lambda_d = 1/4$) i.e. ΔG increased even though pedigree information was removed. The corresponding ratios for ΔF were 0·71 and 0·72, indicating that the incorporation of progeny testing made the use of I_1 more effective in reducing ΔF whilst maintaining progress.

Index I_2. The results of this strategy with and without progeny testing are shown in Table 4. In the absence of progeny testing the proportional inflation of $\sigma_{a,0}^2$, by 0·25 resulted in 0·99 of the response obtained with no inflation (i.e. the basic scheme) with ΔF being 0·90 of that for the basic scheme whereas the larger proportional bias of 0·75 resulted in 1·09 times the response (i.e. increased response) with ΔF being only 0·71 of that for the basic schemes. With progeny testing, when compared with no inflation of $\sigma_{a,0}^2$ both 0·25 and 0·75 inflation of $\sigma_{a,0}^2$ gave a proportional increase of 1·09 in ΔG and a proportional decrease in ΔF of 0·93 and 0·71 respectively. Inflating $\sigma_{a,0}^2$ increased generation intervals in both sexes in the absence of progeny testing, whilst with progeny testing generation intervals were increased for females but decreased for males which resulted in an overall decrease.

Index I_3. For I_3, ΔG was 0·1062 and 0·1316 with and without progeny testing, with ΔF = 0·0091 and 0·0035 respectively. Proportionately this was 0·85 of ΔG and 0·34 of ΔF obtained in the basic scheme without progeny testing and, correspondingly, 1·02 and 0·29 with progeny testing. Without progeny testing a generation interval of 3·85 (s.e. ≤0·06) years was observed in both sexes, greater than in the basic BLUP scheme. With progeny testing generation intervals were 6·18 (s.e. 0·06) and 2·67 (s.e. 0·01) years for males and females, giving an overall interval of 4·43 years.

Discussion

The application of MOET in genetic improvement has been widely evaluated as a strategy in dairy cattle breeding to improve genetic progress, but it is the disadvantageous aspects of using this modern reproductive technology that has attracted increasing attention from animal breeders and quantitative geneticists. In terms of quantitative genetics, an effective breeding scheme should not only make genetic response but also control inbreeding. To satisfy medium or long-term goals it may be necessary to sacrifice some of the short-term genetic gain in order to satisfy a constraint on effective population size. To minimize such sacrifice requires identification of aspects of indices used for selection and breeding practices that promote lineal proliferation without significant contribution to progress. The present study mainly focuses on investigating these possibilities in dairy MOET nucleus schemes and including characteristics inadequately modelled or less studied to date.

Firstly the generator of embryo yields incorporated fully the variation and repeatability observed (Lohuis *et al.*, 1993; Woolliams *et al.*, 1994) in practice (CV = 1·2, r^2 = 0·23). This variation is important in the present context since variation in family size in the short or long term is the driving force behind inbreeding processes and the rate of inbreeding is expected to vary with the square of the coefficient of variation of physical family size (Robertson, 1961; J. A. Woolliams, unpublished). Previous models of embryo yield or calf numbers have included constants (Ruane and Thompson, 1991), scaled binomials (Colleau, 1992) and poisson with zeros (Schrooten and van Arendonk, 1992); none of these distributions includes any repeatability and if the latter two were used to provide the same mean number of embryos per recovery with the same proportion of null recoveries then the CV obtained would be 0·65 and 0·85 respectively.

The work presented has also concentrated on overlapping generations with the use of BLUP for evaluation and selection based, in some form, on the resulting breeding values. Overlapping generations were chosen not only because they are used in

practice and utilize one of the envisaged benefits of BLUP, namely selection across age groups (Hill and Meyer, 1988), but also because modification of the generation interval gives an important additional degree of freedom for changing structures and outcomes; it affects genetic progress proportional to L^{-1} but inbreeding proportional to L^{-2}. Schemes with only moderate nucleus sizes were used which resulted in high rates of inbreeding. Whilst such rates of inbreeding in practice could be alleviated by promoting a wider genetic base, there are potential advantages of scale for this type of study, both computationally and for identifying trends. Of the strategies considered the progeny testing and repeated use of donors are particular to overlapping generations where an individual has several opportunities to breed, whereas the others involving modification of statistical procedures for evaluation and selection or mating design and family sizes are also applicable to discrete generations.

Of all the strategies examined in this study the most effective single measure to control inbreeding and its associated risk was the organized progeny testing of young bulls. In dairy schemes the gain in accuracy from a progeny test is, even in the short term, nearly sufficient to offset the longer generation interval required but the annual rate of inbreeding is however much reduced since as mentioned previously, it is related to L^{-2}. Whilst the results are superficially similar to those of Bondoc and Smith (1993), this study (unlike the previous one) simulated a progeny testing scheme that produced comparable rates of progress to other schemes. This was achieved by conducting the progeny tests outside the nucleus which is also the likely case in practice. Furthermore the deterministic model of Woolliams and Meuwissen (1993) and Meuwissen and Woolliams (1994) showed that in the schemes actively constrained by variance of response progeny testing is essential. There is a close relationship between risks defined by inbreeding and variance of response (Hill, 1979). Therefore it is apparent that organized, effective progeny testing schemes are still a requirement even when much of the breeding activity is conducted in the nucleus.

Progeny testing also gave added potency to the index modifications involving the subtraction of parental breeding values (i.e. I_1). In the absence of progeny testing, inflation of the assumed genetic variance in the base (I_3) was more effective in protecting genetic progress and reducing inbreeding than index I_1. With the inclusion of progeny testing, which was beneficial in itself, these two index modifications appeared to have broadly similar benefits. Explanations for this can be advanced. The estimated breeding value for an individual from BLUP can be decomposed into a finite weighted sum of estimates of the Mendelian sampling term of the individual and all his ancestors, the weighting given to an ancestor t generation back being 2^{-t} (Woolliams and Thompson, 1994). To reduce inbreeding some reduction in the accuracy of selection will be required. The index I_1, proposed by Verrier *et al.* (1993), sacrifices this information uniformly over all generations of ancestors from $t = 1$ (parents) onwards. In contrast a first-order approximation (R. Thompson, unpublished) of the effect of inflating the additive variance is to sacrifice the information progressively with t according to a geometric series (i.e. s^t where $s < 1/2$). Therefore, since individual males have little or no information upon their own Mendelian sampling term without progeny testing, the reweighting of information possible through the use of I_1 is very restricted. This is in contrast to inflating the genetic variance, when substantial re-weighting is still possible even in the absence of information on the candidate's Mendelian sampling term.

The analysis of the re-weighting of the index information described in the previous paragraph can lead to the hypothesis, by an inductive argument, that the components generating least inbreeding for a given increase in accuracy are those of most recent origin. Therefore, with this hypothesis, we might expect to observe that inflating the genetic variance would prove more effective even with progeny testing. Whilst the data from this study show trends supporting this hypothesis the evidence is inconclusive. The results of Villanueva, Woolliams and Simm (1994) in their study of a beef nucleus scheme provide much stronger evidence on this point in a different context.

The explanation for the effectiveness of penalizing relatively more inbred individuals is more difficult since the future rate of inbreeding is determined more by the group and their interrelationship rather than a single individual and its own inbreeding coefficient. In this application its effectiveness may have more to do with its effect on lengthening the generation interval since older individuals will tend to be the less inbred and consequently favoured.

The benefits of factorial mating designs were not as great in this study compared with other studies (e.g. de Boer and van Arendonk, 1994). The reason for this may lie in the large variation in embryo recovery used in simulations. The benefits of factorial mating lie in the removal of full-sibs, since with full-sib families the success or otherwise of one parent affects fewer mates where in factorial mating this fate is more equally shared among all possible mates. With increasing variation in recovery a single group of

full-sibs could dominate the family of a single female. In this way the benefits of factorial mating are diminished.

Reducing the expected family size and increasing the number of donors has direct effect upon the potential size of a donor's expected long-term contribution (equivalently reducing the variance of family size). Therefore the reduction in inbreeding was clear, whilst the reduction in genetic progress was only minor. In schemes of this size this was not unexpected, since Meuwissen (1991) has shown that the effect of small numbers and correlated indices made the true selection intensity achieved insensitive to the numbers selected. One of the options considered was to restrict the opportunities of using donors over more than one period; this can be regarded as essentially the same as using more donors and thereby restricting expected family size.

Toro and Perez-Enciso (1990) investigated various strategies to restrict inbreeding below a fixed level without a significant loss of genetic response. This was obtained mainly through minimizing coancestry of matings. Methodologically, their use of integer linear programming provided a flexible solution for arranging matings with minimal coancestry but is computationally heavy. This may offer an alternative for mating designs in MOET schemes although there are yet further alternatives such as that of Caballero and Santiago (1994) where matings are decided on the basis of sibship sizes post-selection.

In conclusion the results of the present study demonstrated that a large proportion of the inbreeding rate in standard MOET nucleus schemes can be removed without causing loss in genetic response. In the present simulation model, inbreeding was only taken into account as a factor influencing the dynamics of genetic variance although inbreeding might also affect individuals through depression in performance and reduced survival (Wiener, Lee and Woolliams, 1992) through non-additive gene action. Taking inbreeding depression into account would lead to further benefits in genetic response from a reduction in the rate of inbreeding.

Acknowledgement

We thank Drs G. Simm and B. J. McGuirk for their constructive comments on this paper and for their helpful discussions while the study was carried out. Two anonymous referees are acknowledged for their constructive criticisms which were very helpful in improving the presentation of the paper. We are grateful to the Ministry of Agriculture, Fisheries and Food (UK), the Milk Marketing Board of England and Wales and the Meat and Livestock Commission for funding this project.

References

Boer, I. J. M. de and Arendonk, J. A. M. van. 1994. Additive response to selection adjusted for effects of inbreeding in a closed dairy cattle nucleus assuming a large number of gametes per female. *Animal Production* **58:** 173-180.

Bondoc, O. L. and Smith, C. 1993. Optimized testing schemes using nucleus progeny, adult MOET siblings, or juvenile MOET pedigrees in dairy cattle closed populations. *Animal Breeding and Genetics* **110:** 30-40.

Cabellero, A. and Santiago, E. 1994. Prediction of effective size of populations under selection and non-random mating. *Proceeding of the fifth world congress on genetics applied to livestock production, vol. 19,* pp. 143-146.

Colleau, J. J. 1985. [Genetic improvement by ET within selection nuclei in dairy cattle.] *Genetics Selection Evolution* **17:** 499-538.

Colleau, J. J. 1992. Combining the use of embryo sexing and cloning within mixed MOET for selection on dairy cattle. *Genetics Selection Evolution* **17:** 345-361.

Grundy, B. and Hill, W. G. 1993. A method for reducing inbreeding with best linear unbiased prediction. *Animal Production* **56:** 427 (abstr.).

Hill, W. G. 1979. A note on effective population size of populations with overlapping generations. *Genetics* **92:** 317-322.

Hill, W. G. and Meyer, K. 1988. Developments in methods for breeding value and parameter estimation in livestock. In *Animal breeding opportunities occasional publication, British Society of Animal Production, no. 12,* pp. 81-95.

Juga, J. and Maki-Tanila, A. 1987. Genetic change in a nucleus breeding dairy herd using embryo transfer. *Acta Agriculturae Scandinavia* **37:** 511-519.

Leitch, H. W., Smith, C., Burnside, E. B. and Quinton, M. 1995. Effects of female reproductive rate and mating design on genetic responses in closed nucleus schemes. *Animal Science* In press.

Lohuis, M. M., Smith, C. and Dekkers, J. C. M. 1993. MOET results from a dispersed hybrid nucleus programme in dairy cattle. *Animal Production* **57:** 369-378.

Luo, Z. W., Woolliams, J. A. and Simm, G. 1994. Assessment of present and future effectiveness of embryological techniques. *AgBiotech News and Information* **6:** 13N-18N.

Meuwissen, T. H. E. 1991. Expectation and variance of genetic gain in open and closed nucleus and progeny testing schemes. *Animal Production* **53:** 133-141.

Meuwissen, T. H. E. and Woolliams, J. A. 1994. Maximizing genetic response in breeding schemes of dairy cattle with constraints on variance of response. *Journal of Dairy Science* **77:** 1905-1916.

Nicholas, F. W. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. *Animal Production* **36:** 341-353.

Quaas, R. L. and Pollak, E. J. 1980. Mixed model methodology for farm and ranch beef cattle testing program. *Journal of Animal Science* **51:** 1277-1287.

Robertson, A. 1961. Inbreeding in artificial selection programmes. *Genetical Research* **2:** 189-194.

Ruane, J. and Thompson, R. 1991. Comparison of simulated and theoretical results in adult MOET nucleus schemes for daily cattle. *Livestock Production Science* **28:** 1-20.

Schrooten, C. and Arendonk, J. A. M. van. 1992. Stochastic simulation of dairy cattle breeding schemes: genetic evaluation of nucleus size and type. *Journal of Animal Breeding and Genetics* **109:** 1-15.

Strandén, I., Maki-Tanila, A. and Mäntysaari, E. A. 1991. Genetic progress and rate of inbreeding in a closed adult MOET nucleus under different mating strategies and heritabilities. *Journal of Animal Breeding and Genetics* **108:** 401-441.

Toro, M. and Perez-Enciso, M. 1990. Optimization of selection response under restricted inbreeding. *Genetics Selection Evolution* **22:** 93-107.

Verrier, E., Colleau, J. J., Foulley, J. L. 1993. Long term effects of selection based on the animal model BLUP in a finite population. *Theoretical and Applied Genetics* **87:** 446-454.

Villanueva, B., Woolliams, J. A. and Simm, G. 1994. Strategies for controlling rates of inbreeding in adult MOET nucleus schemes for beef cattle. *Genetics Selection Evolution* In press.

Wiener, G., Lee, G. J. and Woolliams, J. A. 1992. Effects of rapid inbreeding and of crossing inbred lines on conception rate, prolificacy and ewe survival in sheep. *Animal Production* **55:** 115-121.

Wiggans, G. R., Misztal, I. and Van Vleck, L. D. 1988. Implementation of an animal model for genetic evaluation of dairy cattle in United States. *Journal of Dairy Science* **71:** *suppl. 2*, pp. 54-69.

Woolliams, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49:** 1-14.

Woolliams, J. A., Luo, Z. W., Villanueva, B., Waddington, D., Broadbent, P. J., McKelvey, W. A. C. and Robinson, J. J. 1994. Analysis of factors affecting embryo yields and ovulation rates. *Journal of Agricultural Science, Cambridge* In press.

Woolliams, J. A. and Meuwissen, T. H. E. 1993. Decision rules and variance of response in breeding schemes. *Animal Production* **56:** 179-186.

Woolliams, J. A. and Thompson, R. 1994. A theory of contributions. *Proceeding of the fifth world congress on genetics applied to livestock production, vol. 19*, pp. 127-134.

Woolliams, J. A., Wray, N. R. and Thompson, R. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research* **62:** 231-242.

Wray, N. R. and Thompson, R. 1990. Prediction of rates of inbreeding in selected populations. *Genetical Research* **55:** 41-54.

(*Received* 9 *February* 1994—*Accepted* 27 *November* 1994)

Paper 24

Table 1 *Number of lambs weaned and total litter weight at weaning (kg) per ewe mated, according to inbreeding (F)*

Inbreeding of ewe (F)	No. of records	No. of lambs weaned per ewes mated		Total litter weight at weaning per ewe mated (kg)	
		No.	s.e.	Weight	s.e.
0·00 (LC)	640	1·42	0·046	34·5	1·00
0·00 (O_2/F_2)	679	1·23	0·060	27·5	1·29
0·25	842	1·00	0·036	20·7	0·78
0·375	958	0·75	0·034	15·2	0·73
0·50	795	0·63	0·039	12·4	0·83
0·59	337	0·76	0·058	16·4	1·26

For purposes of the present study the following criteria have been applied.

1. Age at first lambing: 2 years, and age at disposal of surviving ewes in the breeding flock: 5·5 years, after four potential lamb crops (as actually practised).
2. Replacement females kept only in numbers adequate to maintain a flock in regular ages and of constant size.
3. Surplus female lambs and all male lambs sold for slaughter at the age of 27 weeks (12 weeks after weaning) — the earliest target age in this flock.

No reduction in numbers is made for any male lambs which might be required as breeding rams. Many lambs, especially the more highly inbred, took longer than 27 weeks to reach an acceptable condition for slaughter. However, experience on the farm in question showed that, for this experimental flock, the extra money obtained from selling lambs later than 27 weeks old, at the heavier weights and improved condition then reached, was offset by the extra costs involved. Thus the use of a fixed age at slaughter and a fixed price per kg live weight, for purposes of the calculation of financial returns, is considered to provide a valid comparison among the inbreeding classes.

4. The average prices obtained in 1992 for the class of stock in this flock on this farm have been used and are: per kg live weight of lamb — £0·70; per kg live weight of cast ewes (at the end of breeding life) — £0·45; per kg greasy wool — £0·85.

The starting point for the calculations of financial returns has been an analysis of the number of lambs weaned per ewe mated and the total litter weight at weaning per ewe mated, traits not considered in earlier papers. The statistical model used for these analyses were as described for the analysis of prolificacy (Wiener *et al.*, 1992b). All other data were derived from papers previously published: for survival to 27 weeks (Wiener, Woolliams and Macleod, 1983) survival thereafter (Wiener *et al.*, 1992b), growth rate (Wiener *et al.*, 1992a) and fleece weight (Wiener *et al.*, 1994).

Results and discussion

Inbreeding of ewe

Table 1 shows, according to the inbreeding level of the ewe, the number of lambs weaned and the total litter weight at weaning (male and female lambs combined) per ewe mated. These 'composite' traits comprise several components each of which may be affected by inbreeding. The results show a large, significant reduction ($P < 0·001$) in the number of lambs weaned from 1·42 per LC ewe mated to 0·63 for the F 0·50 class — with a small increase with further inbreeding to F 0·59. An apparent improvement in performance between F 0·50 and F 0·59 was also observed in previously published analyses and possible reasons for the observation discussed, in particular in the study related to growth (Wiener *et al.*, 1992a). The relative decline with inbreeding in litter weight per ewe mated ($P < 0·001$) is seen to be even greater due to the additional adverse effects of inbreeding on lamb growth. Inbreeding to F 0·50 is seen to have more than halved the output of lamb (but, as also noted above, the output from the F 0·59 class was better than that from F 0·50).

Taking into account the reproduction data in Table 1 and survival data on lambs published by Wiener *et al.* (1983) and on ewes by Wiener *et al.* (1992b), it was found that it would need all of the potential four lamb crops to maintain the F 0·50 class as a flock of constant size — though a lesser number of lamb crops for the other inbreeding classes. In consequence, the age structure of a breeding flock of constant size is slightly different for each of the inbred classes (See Table 2). Thus, the proportion of young females relative to older females has to increase with inbreeding.

Anim. Prod. 1994, 59: 245-249

0003-3561/94/3315024550200

Consequences of inbreeding for financial returns from sheep

G. Wiener, G. J. Lee† and J. A. Woolliams

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS

Abstract

The effects of inbreeding on the gross financial returns from ewes have been studied in a flock of sheep comprising three hill breeds and the crosses among them. Each of these groups was represented by non-inbred sheep (purebred and crossbred, designated O_2 and F_2 respectively) and four classes of inbreeding (inbreeding coefficients (F): 0·25, 0·375, 0·50 and 0·59). A second non-inbred class was represented by crosses of inbred lines.

Gross lifetime income per ewe derived from the sale of lambs for slaughter, cast ewes and wool was reduced by inbreeding in an almost linear fashion (between the O_2/F_2 and the F 0·50 class) at a rate of £1·27 (at 1992 prices) for each 0·01 (percentage point) increase in the level of inbreeding — an overall reduction to 0·35 of the O_2/F_2 level. Line crosses gave a better return than the other non-inbred sheep. When expressed in terms of ewe metabolic body weight ($M^{0.73}$) the decline with inbreeding was less — an overall reduction to 0·4 of the O_2/F_2 income. Inbreeding of the lamb caused a further decline in output independent of the inbreeding of the ewe.

Keywords: *inbreeding, returns, sheep.*

Introduction

Many studies have shown that inbreeding reduces the performance in various traits of economic importance in sheep (for a review, see Lamberson and Thomas, 1984). A long-term experiment involving deliberate, close inbreeding of sheep on an upland grazing farm showed deleterious effects of inbreeding on reproductive rate, survival, growth rate and wool production (Wiener, Lee and Woolliams, 1992 a,b and 1994).

In order to assess the various effects of inbreeding on the overall output from sheep production, an attempt is made in this paper to relate the inbreeding effects to gross financial returns.

Material and methods

The data are derived from an experiment described fully by Wiener *et al.* (1992a). Three hill breeds and the reciprocal crosses among them were involved, *viz*, Scottish Blackface, Cheviot and Welsh Mountain. Inbreeding started from a purebred or a crossbred base (designated O_2 for purebred and F_2 for crossbred) and involved mostly younger parent × offspring matings for four generations of both the purebred and the crossbred animals. This scheme resulted in inbreeding coefficients (F) from 0·0 to 0·59. Crosses of inbred lines (LC) were made also.

The experiment was conducted on an upland grazing farm, Blythbank, Tweeddale, Scotland. The sheep were kept as a single flock thus ensuring that sheep of the different breeds, crosses and inbreeding classes had equality of opportunity. No artificial selection was practised. The experiment spanned 17 years. The inbreeding classes overlapped in time and the design of the experiment allowed separation of the effects of inbreeding of the lamb from those of its dam.

The flock size throughout the experiment was around 300 breeding females per year, plus followers, though the composition changed over the years in terms of inbreeding classes. Ewes were mated to lamb for the first time at 2 years old and annually thereafter. Survivors were sold at the age of 5·5 years. For experimental reasons, females surplus to requirements for breeding were not sold until 1·5 years old and many more males were kept for breeding to sustain the inbreeding system than would be normal in a commercial flock.

† To whom reprint requests should be addressed.

Table 2 *Estimated age structure of a flock of 100 breeding females, kept for up to four lamb crops, according to inbreeding (F)*

Inbreeding of ewe (F)	No. of females mated in each age group for lambing at age (years)			
	2	3	4	5
0·00 (LC/O_2/F_2)	26·2	25·5	24·6	23·7
0·25	26·9	25·8	24·4	22·9
0·375	27·4	26·0	24·2	22·4
0·50	28·0	26·3	24·0	21·7
0·59	28·6	26·5	23·8	21·1

Differences in age structure have a further effect on lamb numbers because of a parity effect on litter size, shown previously (Wiener *et al.*, 1992b). Because the different number of female lambs born and surviving per ewe over its lifetime varies for each of the inbreeding classes, the number of female lambs surplus to replacement needs and available for slaughter declines with inbreeding. Thus, LC ewes have a substantial number of surplus females available for slaughter and the F 0·50 class of ewe very few. All males surviving to 27 weeks old are assumed available for slaughter. Table 3 shows the numbers of males and females and the total weight of lamb at slaughter for a flock of each inbreeding class over a 4-year period (lifetime of the ewe). An adjustment has been made for the effect of sex of lamb on the weights.

Table 3 *Estimated effect of inbreeding (F) on the number of male and female lambs available for slaughter at 27 weeks old and their total weight (kg) in a flock of 100 ewes with a potential of four lamb crops*

Inbreeding of ewe (F)	Males		Females	
	No.	Total weight	No.	Total weight
0·00 (LC)	278	8562	172	5043
0·00 (O_2/F_2)	248	7608	142	4157
0·25	194	5562	85	2330
0·375	144	3887	32	819
0·50	121	3097	6	145
0·59	142	3772	24	611

Table 4 *Estimated total weight (kg) of cast ewes sold for slaughter and estimated total weight (kg) of greasy wool clipped per 100 ewes, according to inbreeding (F)*

Inbreeding of ewe (F)	Total weight of ewes slaughtered	Total weight of wool clipped
0·00 (LC)	1299	963
0·00 (O_2/F_2)	1299	1005
0·25	1110	922
0·375	1011	841
0·50	886	788
0·59	935	713

Ewes surviving to 5·5 years of age (following four potential lamb crops) are assumed sold for slaughter according to weight. Weights of adult females varied with inbreeding as previously reported (Wiener *et al.*, 1992a). The weights multiplied by the numbers are shown in Table 4 for each inbreeding class per 100 ewes in the breeding flock. Also shown in Table 4, per 100 ewes in the breeding flock, is the total weight of greasy wool clipped. However, there was an additional shearing at 1·3 years of age — wool from young replacement females has therefore been included.

Table 5 shows the contribution of the separate components of output to gross income per 100 ewes, using the prices shown earlier. It can be seen from Table 5 that, as is normal, the major part of the income is derived from the sale of slaughter lambs. Total gross income, over the ewe's lifetime, declined rapidly with an increasing level of inbreeding at a rate, seen to be close to linear, of £1·27 (at 1992 prices) per ewe for each 0·01 (percentage point) of inbreeding over the range from the O_2/F_2 stage to F 0·50. Crossing of inbred lines provided a clear benefit, above the level achieved by the O_2/F_2 class,

Table 5 *Gross lifetime income (£) from different sources, per 100 ewes, according to inbreeding (F) of ewe*

Inbreeding of ewe (F)	Source of gross income (£)					Reduction in income (£) relative to LC
	Male lambs	Female lambs	Cast ewes	Wool	Total	
0·00 (LC)	5994	3530	585	819	10928	
0·00 (O_2/F_2)	5326	2910	585	854	9675	1253
0·25	3893	1631	499	784	6807	4121
0·375	2721	574	455	715	4465	6463
0·50	2168	102	399	670	3339	7589
0·59	2640	428	421	606	4095	6833

as a result of an accumulation of several small advantages in component traits. The gross income from the F 0·59 class of ewes exceeded that from their F 0·50 contemporaries in line with a similar trend reported previously for some of the component traits.

Inbreeding of lamb

In the earlier studies referred to it was found that inbreeding of lamb had a further effect on performance independent of the effect of inbreeding of the dam. This should be taken into account in estimating gross financial return. The estimates in Table 5 have assumed an average type of lamb from each class of ewe. In the present experiment each of the inbred types of ewe generated inbred offspring only more inbred than themselves unless mated to generate line crosses. However, if inbreeding occurs at only a slow rate, as might happen in a closed commercial flock, the inbreeding of the lamb may be very similar to that of the ewe (just slightly higher). The LC ewes themselves can be assumed to generate further LC lambs and the O_2/F_2 class to generate either non-inbred or inbred lambs.

Table 6 provides estimates of the further reductions in income from lowered slaughter weights of lamb resulting from inbreeding of the lamb. It thus appears, from Table 6, that gross income (per 100 ewes) can be further reduced by inbreeding of the lamb by the amounts shown in the second column of the Table. The figures combine the increasing reduction in income per lamb with the decreasing number of lambs of each type sold. This further reduction in income from the ewe due to inbreeding of the lamb represents about one-fifth of the gross income shown for each of the three highest inbreeding classes of ewe in the second last column of Table 5.

Age structure

Because increasing inbreeding involved an age structure of ewe flock with a larger proportion of young females, some extra costs are involved in rearing these extra females for a year after they might otherwise have been available for slaughter. For the farm in question this is estimated at £15 per female and relative to the two non-inbred classes involves rearing costs per 100 ewes of (£) 11, 18, 27 and 36 for the four inbred classes (F 0·25, 0·375, 0·50 and 0·59) respectively. Contrary to some of the other findings for the F 0·59 class of ewe, they are here at a disadvantage to the F 0·50 class because mortality continued to increase with inbreeding.

Effects of body weight of the dam

Webster (1989) has pointed out that proportionately 0·70 of the total metabolizable energy required for lamb production (to slaughter) is consumed by the ewe. It is often assumed that maintenance costs of small dams are lower than those of heavier ones due to lower food intake. These assumptions were made, in relation to the present experimental flock, by Wiener (1967) who pointed out that per kg $M^{0.73}$ the output from the small Welsh Mountain breed exceeded that from some of the larger breeds involved. A similar calculation was made by Lee, Haley and Land (1991) in assessing the merits of a breeding scheme intended to change prolificacy but which also affected body size of the ewe.

These considerations are relevant here because adult size of the females in this experiment declined with inbreeding. Some support for the idea that this might provide an advantage in terms of maintenance costs per ewe was provided by a study of Wiener, Woolliams and Slee (1988). It showed that in order to maintain live weight, inbred sheep (F 0·375 and 0·50) consumed, on an indoor feeding regime, the same

Table 6 *Effect of inbreeding (F) of lambs on gross income*

Inbreeding of lamb (F)	Reduction in income (£) per lamb relative to LC lamb	Further reduction† in lifetime income per 100 ewes if inbreeding coefficient of the ewe is similar to that of the lamb‡ (£)
0·00 (LC)		
0·00 (O_2/F_2)	1·12	111
0·25	3·47	1002
0·375	5·05	944
0·50	5·08	716
0·59	4·45	819

† This 'further' reduction in income may be added to the reduction shown in the last column of Table 5 attributable to inbreeding of the ewe.

‡ This would assume a slow rate of further inbreeding.

Table 7 *Gross income per kg metabolic live weight ($M^{0.73}$) of ewe, according to inbreeding (F)*

Inbreeding of ewe (F)	Gross income (£) per kg $^{0.73}$ of ewe mated
0·00 (LC)	5·79
0·00 (O_2/F_2)	5·12
0·25	3·91
0·375	2·69
0·50	2·14
0·59	2·46

amount of food per kg $M^{0.73}$ live weight as did, otherwise, similar, non-inbred sheep.

Table 7 shows the gross income per kg $M^{0.73}$ of ewe for each of the inbreeding classes. It shows that gross income declined less rapidly with increased inbreeding if calculated on the basis of metabolic live weight than when calculated on a per ewe basis (cf. Table 5).

Conclusions

In conclusion it can be seen that inbreeding to F 0·50 reduced gross income from the flock very substantially — by a factor of 2·5 to 3 compared with a non-inbred base, and only a little less if account is taken of the effect of inbreeding on ewe size.

The single largest contribution to the financial returns came from the sale of lambs for slaughter. The effects of inbreeding on each of the component traits of overall productivity have been shown to be statistically significant (generally, $P < 0.001$) in the present (Table 1) and the previous studies referred to which have been used as the sources of data. That combined evidence provides some confidence in the importance of the effects of inbreeding of the ewe on the financial returns (despite the absence of a direct statistical test).

Further reductions in income were due to inbreeding of the lamb and due to greater replacement costs with increased inbreeding. The levels of inbreeding reached in the present experimental flock were higher than those which might be expected in most commercially run flocks. However, it can be inferred that even small amounts of inbreeding, as might occur in commercial breeding flocks, can be deleterious for income. This is likely because the effects of inbreeding (up to F 0·50) on financial returns shown here were largely linear and because for the component traits, in previously published studies, the changes per degree of inbreeding were similar to those reported by Lamberson and Thomas (1984) who reviewed a large number of studies involving mostly much lower levels of inbreeding.

Where returns from sheep are affected by subsidy payments per ewe (or lamb) as, for example, in the EC, the relative disadvantage of inbreeding is reduced — but this is a non-biological distortion to the effects of inbreeding which is not considered here.

Acknowledgements

We are grateful to Mr David Drury and Mr Phil Davies of Blythbank farm for useful discussions concerning the management of the flock and marketing of the output.

References

Lamberson, W. R. and Thomas, D. L. 1984. Effects of inbreed in sheep: a review. *Animal Breeding Abstracts* **52:** 287-297.

Lee, G. J., Haley, C. S. and Land, the late R. B. 1991. Correlated changes in growth patterns and overall production efficiency following selection for testis size adjusted for body weight in young male lambs. *Animal Production* **52:** 477-488.

Webster, A. J. F. 1989. Bioenergetics, bioengineering and growth. *Animal Production* **48:** 249-269.

Wiener, G. 1967. A comparison of the body size, fleece weight and maternal performance of five breeds of sheep kept in one environment. *Animal Production* **9:** 177-195

Wiener, G., Lee, G. J. and Woolliams, J. A. 1992a. Effects of rapid inbreeding and of crossing of inbred lines on the body weight of sheep. *Animal Production* **55:** 89-99.

Wiener, G., Lee, G. J. and Woolliams, J. A. 1992b. Effects of rapid inbreeding and of crossing of inbred lines on conception rate, prolificacy and ewe survival in sheep. *Animal Production* **55:** 115-121.

Wiener, G., Lee, G. J. and Woolliams, J. A. 1994. Effects of breed, rapid inbreeding, crossbreeding and environmental factors on fleece weight and fleece shedding in sheep. *Animal Production* **59:** 61-70.

Wiener, G., Woolliams, C. and Macleod, N. S. M. 1983. The effects of breed, breeding system and other factors in lamb mortality. 1. Causes of death and effects on the incidence of losses. *Journal of Agricultural Science, Cambridge* **100:** 539-551.

Wiener, G., Woolliams, C. and Slee, J. 1988. A comparison of inbred and outbred sheep on two planes of nutrition. 1. Growth, food intake and wool growth. *Animal Production* **46:** 213-220.

(*Received* 6 *September* 1993—*Acccepted* 22 *April* 1994)

Paper 25

Animal Science 1996, 63: 563-576

1357-7298/96/57710563$20-00

Optimum designs for breeding programmes under mass selection with an application in fish breeding

B. Villanueva[1], J. A. Woolliams[2] and B. Gjerde[3]

[1]*Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG*
[2]*Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS*
[3]*AKVAFORSK, PO Box 5010, N-1432 Ås, Norway*

Abstract

A procedure for maximizing genetic gain (after a number of generations of selection) for a given rate of inbreeding or for a given coefficient of variation of response is presented. An infinitesimal genetic model is assumed. Mass selection is practised for a number of discrete generations. With constraints on inbreeding, expected rates of genetic progress (ΔG) are combined with expeced rates of inbreeding (ΔF) in a linear objective function ($\Phi = \Delta G - \lambda \Delta F$). In addition, an expression to approximate the rate of gain at any generation accounting for changes in genetic parameters due to linkage disequilibrium and due to inbreeding is derived. Predicted gain is in general within 5% of that obtained from simulation. Thus, both ΔG and ΔF are obtained from simple analytical formulae. An equivalent function is used when the coefficient of variation of response (CV) is the parameter restricted ($\Phi = \Delta G - \lambda CV$). Maximization of the objective function Φ for appropriate values of λ gives the optimum number of sires and dams selected when specific constraints on the level of inbreeding or the coefficient of variation of response are imposed. The method is applied to a practical situation in fish breeding. Optimum mating ratios and optimum numbers of sires selected are obtained for different scored population sizes and heritabilities. Results obtained with this procedure agree very well with results from simulation studies. The optimum number of sires increases with the size of the scheme and with more severe restrictions on risk. In the schemes considered, the optimum mating ratio is equal to 2 unless the constraint on the rate of inbreeding is severe, the size of the scheme is small and the heritability is low. In these situations the optimum mating ratio is equal to 1. The procedure is general in terms of generations of selection considered and in terms of parameters to be constrained. A large amount of computer processor unit time is saved with this method in comparison with simulation procedures.

Keywords: *breeding programmes, fish, genetic gain, inbreeding.*

Introduction

Modern techniques for selective breeding are applied in most agricultural species and their use has considerably altered the performance of farmed animals. However, there is now a growing concern about the long-term consequences of the loss of genetic variability as a result of selection. Considerable attention has been paid recently to selection and mating procedures for controlling inbreeding to reduce its detrimental effects (loss of genetic variation and inbreeding depression). An alternative but related measure of the risk of breeding schemes, the variability of selection response, has been also considered by several authors (Goddard, 1987; Nicholas, 1989; Meuwissen and Woolliams, 1994a).

Fish species have not fully benefited from modern developments in animal breeding theory despite their high potential for genetic improvement through selective breeding (Bentsen and Gjerde, 1994). When compared with other farmed animals, fish have a very high reproductive capacity which allows high selection intensities. However, high inbreeding is also expected if a small number of individuals contribute largely to subsequent generations. The problem of inbreeding can be important even with selection procedures which do not make use of family information; i.e. mass selection (e.g. Bentsen and Gjerde, 1994).

Optimum designs for mass selection programmes in fish breeding under constrained inbreeding have

been investigated by Gjerde *et al.* (1996) through stochastic simulation. For a given number of scored individuals and mating ratio, rates of gain and inbreeding were obtained separately for different numbers of sires. Designs giving a rate of inbreeding reasonably close to a specific value were defined as optimum.

More general and less computationally demanding results could be obtained from analytical methods but the problem of predicting inbreeding in selected populations has not been solved until recently. Now, accurately and tractable expressions for the expected rate of inbreeding in populations under mass (Woolliams *et al.*, 1993; Woolliams and Thompson, 1994; Santiago and Caballero, 1995) and index selection (Wray *et al.*, 1994) are available. Expected rates of inbreeding can be combined with expected genetic gains to obtain optimum solutions for the design of breeding programmes.

In this paper, a procedure for maximizing response while restricted inbreeding is presented. A simple expression for predicting the rate of gain over a variable number of generations and which accounts for the effects of inbreeding is also developed. Optimum fish breeding schemes are obtained by maximizing a single, closed form, objective function which combines expected rates of progress after a number of generations of mass selection with expected rates of inbreeding. An equivalent procedure is used for optimizing schemes while restricting the coefficient of variation of response. Specific restrictions on rate of inbreeding or on variability of response are considered.

Methods

Model

The trait under selection is assumed to be determined by an infinite number of additive loci, each with infinitesimal effect (infinitesimal model). The selection criterion is the phenotypic performance of the candidate for selection. Repeated cycles of selection are practised in discrete generations. Selection is directional and by truncation. Each generation, N_s males and N_d females are selected from $N/2$ males and $N/2$ females scored, and the values of N_s and N_d are optimized for each breeding scheme as described later. Male and female selection intensities are kept constant across generations. Selected individuals are mated at random under a hierarchical design and dams are nested within sires. Each pair of parents produces the same number of male and female offspring.

The total number of individuals scored is considered to be the main constraint of the scheme. This number (N) is fixed each generation. Different values for N and for the initial heritability ($h^2_{(0)}$) are considered.

Prediction of rate of genetic progress

(a) Accounting for reduction in genetic variance due to linkage disequilibrium

Prediction A (P_A). Under the infinitesimal model, if the effects of inbreeding are ignored then the only cause of change in genetic parameters is linkage disequilibrium (Bulmer, 1971). In this situation, genetic response approaches an equilibrium after a few generations of selection. Asymptotic response under mass selection can be obtained from

$$\Delta G_{(L)} = ih^2_{(L)}\sigma_{P(L)}$$

where $i = (0{\cdot}5)(i_m + i_f)$ and i_m and i_f are the selection intensities for males and females, respectively, and $h^2_{(L)}$ and $\sigma_{P(L)}$ represent the limiting or equilibrium values for the heritability and phenotypic standard deviation. Limiting values for the heritability and the phenotypic variance can be obtained from

$$h^2_{(L)} = \frac{-1 + \sqrt{1 + 4h^2_{(0)}k(1 - h^2_{(0)})}}{2k(1 - h^2_{(0)})}$$

and $\sigma^2_{P(L)} = \sigma^2_{A(L)} + \sigma^2_E$
where $\sigma^2_{A(L)} = \sigma^2_{A(0)}/(1 + kh^2_{(L)})$
$k = (0{\cdot}5)(k_m + k_f)$
$k_y = i_y(i_y - x_y)$, for $y = m$ (males) or f (females).

x_y is the standardized deviation of the truncation point from the mean, σ^2_E is the environmental variance (assumed constant over generations), $\sigma^2_{A(0)}$ and $h^2_{(0)}$ are the initial values for the genetic variance and the heritability, respectively and $\sigma^2_{A(L)}$ is the limiting value for the genetic variance (Bulmer, 1971; Gomez-Raya and Burnside, 1990).

(b) Accounting for reduction in genetic variance due to linkage disequilibrium and inbreeding

Two procedures are considered to predict genetic gain when changes in genetic parameters due to linkage disequilibrium and due to inbreeding are both accounted for.

Prediction B (P_B). Let $G_{(t)}$ be the average genetic mean of individuals born at generation t. The rate of genetic gain obtained each generation ($\Delta G_{(t)} = G_{(t)} - G_{(t-1)}$) is

$$\Delta G_{(t)} = ih^2_{(t-1)}\sigma_{P(t-1)}.$$

The heritability and phenotypic variance at generation t are given by

$$h^2_{(t)} = \sigma^2_{A(t)}/\sigma^2_{P(t)}$$

and

$$\sigma^2_{P(t)} = \sigma^2_{A(t)} + \sigma^2_E.$$

The genetic variance is decomposed into between sire family ($\sigma^2_{As(t)}$), between dam family ($\sigma^2_{Ad(t)}$) and within full-sib family ($\sigma^2_{Aw(t)}$) components ($\sigma^2_{A(t)} = \sigma^2_{As(t)} + \sigma^2_{Ad(t)} + \sigma^2_{Aw(t)}$) and these are obtained each generation by using the following recurrent equations:

$$\sigma^2_{As(t)} = (0{\cdot}25)[1 - (1/N_s)][1 - k_m h^2_{(t-1)}]\sigma^2_{A(t-1)}$$

$$\sigma^2_{Ad(t)} = (0{\cdot}25)[1 - (1/N_d)][1 - k_f h^2_{(t-1)}]\sigma^2_{A(t-1)}$$

$$\sigma^2_{Aw(t)} = (0{\cdot}5)[1 - F_{(t-1)}]\sigma^2_{A(0)}$$

where $F_{(t)}$ is the average inbreeding coefficient at generation t (e.g. Verrier *et al.*, 1990). After generation 0, the average inbreeding coefficient is computed as

$$F_{(t)} = 1 - (1 - \Delta F)^{t-1}.$$

The asymptotic rate of inbreeding (ΔF) given by Woolliams and Thompson (1994) is used to obtain $F_{(t)}$ at each generation (see below).

Prediction C (P_C). A closed form was intended to predict genetic gain after the Bulmer equilibrium has been approached. The rate of genetic response at any generation t can be approximated as

$$\Delta G_{(t)} \approx \frac{h^2_{(0)} i (a^{t+1} - b^{t+1})}{2\sigma_{P(L)} a^3 (a - b)} \qquad (1)$$

where $b = (0{\cdot}5)(1 - kh^2_{(L)})$

$$a = \frac{1 - \Delta F}{1 - [(0{\cdot}25)h^2_{(0)}\Delta F]/\sigma^2_{P(L)}}$$

and ΔF is the asymptotic rate of inbreeding which is obtained following Woolliams and Thompson (1994). The derivation for (1) is described in the **Appendix.**

The average rate of gain from generation $t = n$ to generation $t = n + m$ is

$$\Delta\bar{G}_{(n,\,n+m)} = \frac{\Delta G_{(n)} + \ldots + \Delta G_{(n+m)}}{m+1}$$

$$\approx \frac{h^2_{(0)} i}{2\sigma_{P(L)} a^3 (a-b)(m+1)} \left[a^{n+1} \sum_{j=0}^{m} a^j - b^{n+1} \sum_{j=0}^{m} b^j\right]$$

$$\approx \frac{h^2_{(0)} i}{2\sigma_{P(L)} a^3 (a-b)(m+1)} \left[a^{n+1} \frac{1 - a^{m+1}}{1 - a} - b^{n+1} \frac{1 - b^{m+1}}{1 - b}\right] \qquad (2).$$

Prediction of rate of inbreeding

The asymptotic rate of inbreeding (ΔF) was calculated using the approximated expression given by Woolliams and Thompson (1994) corrected for hypergeometric sampling:

$$\Delta F = \frac{1 + i^2 S^2_\infty \rho_m}{16N_s} + \frac{1 + i^2 S^2_\infty \rho_f}{16N_d} + \frac{i^2 S^2_\infty (\rho_m + \rho_f)}{16N_d}$$

$$+ \left[\frac{1}{16N_s} + \frac{1}{16N_d} - \frac{1}{8T}\right]\left[1 - \frac{1}{2N_s} - \frac{1}{2N_d}\right]$$

$$+ i^2 B_\infty (S_\infty - 1)\left[2\rho_m + \left(1 + \frac{N_s}{N_d}\right)\rho_f\right]\left[\frac{1}{32N_s} + \frac{1}{32N_d}\right]$$

$$+ i^2\left[\frac{\rho}{16N_s} + \frac{\rho_f}{16N_d} + \frac{\rho_m + \rho_f}{16N_d}\right]$$

where

$B_\infty = 1/[1 - (0{\cdot}5)c^2]$
$S_\infty = 1/[1 - c]$
$c = (c_m + c_f)/2$
$c_m = (1 - k_m h^2_{(2)})/2$
$c_f = (1 - k_f h^2_{(2)})/2$

$$h^2_{(2)} = h^2_{(0)} \frac{1 - (0{\cdot}5)kh^2_{(0)}}{1 - (0{\cdot}5)kh^4_{(0)}}$$

$$\rho_m = \frac{1}{4} h^2_{(0)} \frac{1 - k_m h^2_{(0)}}{1 - (0{\cdot}5)kh^4_{(0)}}$$

$$\rho_f = \frac{1}{4} h^2_{(0)} \frac{1 - k_f h^2_{(0)}}{1 - (0{\cdot}5)kh^4_{(0)}}$$

and T is the number of selection candidates of each sex ($T = N/2$).

Prediction of variance of genetic progress

The variance of the genetic response after a single generation of selection ($V(\Delta G)$) is taken to be

$$V(\Delta G) \approx 2\Delta F \sigma^2_{A(L)} (1 - h^2_{(L)} k)$$

which was shown by Meuwissen and Woolliams (1994b) to be a good approximation for mass selection. The coefficient of variation of total response between generations n and $n + m$ ($CV_{n,n+m}$) is computed as

$$CV_{(n,n+m)} \approx \sqrt{(m+1)\,V(\Delta G)}/(m+1)\Delta\bar{G}_{(n,\,n+m)}.$$

Optimization of breeding schemes

Optimum designs (with respect to N_s and $d = N_d/N_s$) are obtained by maximizing genetic gain with an upper bound (η) on the rate of inbreeding. For each set of N and $h^2_{(0)}$ values (fixed parameters), the average rate of genetic progress from generation n to generation $n + m$ and the asymptotic rate of inbreeding are obtained as described above. Then, rates of genetic gain and inbreeding are combined in a single objective function (Φ),

$$\Phi = \Delta\bar{G}_{(n,\, n+m)} - \lambda\Delta F$$

which is maximized. For P_A and P_C the Φ is a closed expression determined by $h^2_{(0)}$, N_s and N_d, whereas for P_B the expression is recursive requiring stepwise computation over the $n + m$ generations. The parameter λ takes positive values. From the theory of Lagrangian multipliers, this procedure (to maximize Φ) is equivalent to maximize $\Delta\bar{G}_{(n,\, n+m)}$ subject to the constraint $\Delta F \leq \eta$ as indicated by Woolliams and Thompson (1994). For fixed λ, N and $h^2_{(0)}$, the objective function depends on two variables which are the number of males selected (N_s) and the mating ratio (d). These variables are subject to a non-linear constraint ($N_d/N_s = d$). The problem is to find the value of those variables where Φ takes the maximum value. Alternatively, if required, d can be restricted to some value and then N_s is the only variable to be optimized.

For each possible value of N_s ($1 \leq N_s \leq N/2$), the optimum d was found by using a golden section search in one dimension (e.g. Press *et al.*, 1992). Minimum and maximum values for d are 1 and $N/2N_s$, respectively. The subroutine described in Press *et al.* (1992) was used to obtain the value of d which minimizes $-\Phi$ (this is equivalent to maximizing Φ) for a given N_s. Checks were carried out to establish that the function Φ is unimodal for a fixed value of N_s (i.e. there is guarantee that the identified optimum d is global). As the golden search routine gives a real value for d, the final (integer) value was obtained by searching around the real maximum. The optimum N_s and d were those for which Φ takes the highest value. The number of offspring per mating (n_{off}) is determined by N and N_d ($n_{off} = N/N_d$).

The same approach can be followed for optimizing breeding schemes while constraining the coefficient of variation of response. In this case, the function to be maximized is $\Phi = \Delta\bar{G}_{(n,n+m)} - \lambda CV_{(n,n+m)}$.

The value of λ is varied at appropriate intervals for obtaining rates of inbreeding (or coefficients of variation of response) lower than specific values. Different restrictions on ΔF and $CV_{(n,n+m)}$ were considered. Obviously, when $\lambda = 0$ no constraints are imposed on inbreeding or variability of response.

Computer simulation

The closed form derived here for predicting genetic progress accounting for changes in genetic parameters due to linkage disequilibrium and due to inbreeding (P_C) was tested through stochastic simulation. Genetic values of unrelated base individuals (generation zero) were obtained from a normal distribution with mean zero and variance $h^2_{(0)}$. Phenotypic values were obtained by adding a normally distributed environmental component with mean zero and variance $1 - h^2_{(0)}$. Environmental variance was maintained constant over generations. In each generation (including generation zero), N_s males and N_d females with the highest phenotypic values were selected (from $N/2$ male and $N/2$ female candidates) to be parents of the next generation. Mating of selected individuals was at random and each pair produced N/N_d offspring ($N/2N_d$ males and $(N/2N_d$ females).

True genetic values of the offspring born every generation were generated as the average of the genetic values of their parents plus a random Mendelian deviation. The latter was sampled from a normal distribution with mean zero and variance $(h^2_{(0)}/2)[1 - (F_s + F_d)/2]$, where F_s and F_d are the inbreeding coefficients of the sire and dam, respectively. Inbreeding coefficients were obtained using the algorithm proposed by Meuwissen and Luo (1992). Selection was carried out for 20 generations. One thousand replicates were run for each simulation.

Results

Prediction of genetic progress

Results from using the three predictions for genetic gain described before (P_A, P_B and P_C) were compared with results from stochastic simulations. Figure 1 shows a comparison of simulated values with predictions across generations for a scheme with $N_s = 20$, $N_d = 20$ and $N = 240$ and three values of heritability (0·1, 0·4 and 0·9). In all cases λ was set to zero. For P_A, the genetic gain before the Bulmer equilibrium is approached was obtained by using the recurrent expression given by Gomez-Raya and Burnside (1990). Method C (P_C) was accurate for predicting gain from the third generation onwards, once the Bulmer equilibrium has been approached. Gain is markedly underpredicted in the initial generations because P_C uses equilibrium values for genetic parameters to predict response from the start of selection. Predictions accounting only for reduction in variance due to the Bulmer effect (P_A) considerably overpredict gain. Both P_B and P_C gave very good predictions of genetic gain.

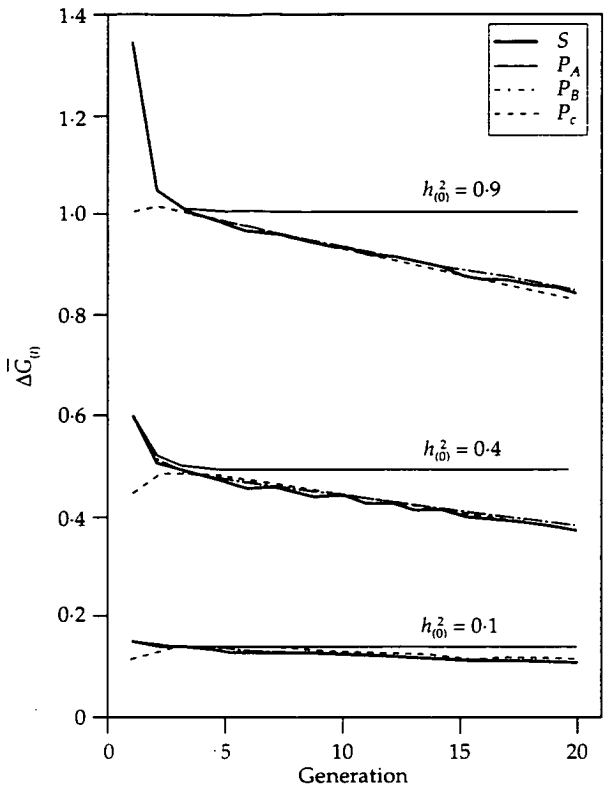


Figure 1 Simulated (S) and predicted (P_A, P_B and P_C) genetic gain ($\sigma_{P(0)}$ units) across generations for three different values of the initial heritability ($h^2_{(0)}$) for a scheme with $N = 240$, $N_s = 20$ and $N_d = 20$.

Table 1 shows simulated values for average rates of gain from generation 5 to generation 14 and percentage errors of predictions (E_{P_A}, E_{P_B} and E_{P_C}). The number of sires (N_s) was 20 in all cases and N_d varied from 20 to 200. Two different family sizes ($n_{off} = N/N_d = 6$ or 12) and a range of initial heritabilities were considered. Standard errors of simulated values for $\Delta G_{(t)}$ ranged from 0·0010 to 0·0034. Accounting only for reduction in genetic parameters due to the Bulmer effect (P_A) led to an overprediction of genetic progress of around 3% to 14%. The bias increases with n_{off} and decreases with N_d, which might be expected since the rate of inbreeding is higher with high n_{off} and low N_d (Woolliams *et al.*, 1993).

Prediction B (P_B) which updates phenotypic and genetic parameters!padh generation and adjusts for finite number of parents had very small errors which were never higher than 2%. In most situations, gain was overpredicted, but for $h^2_{(0)} = 0{\cdot}9$ and high mating ratios, genetic progress was underpredicted. Predictions using method C (P_C) were also very accurate (the highest error was around 5%) although slightly worse than those obtained with method B. Method C uses equilibrium values for heritability and phenotypic variance from the first generation (i.e. different from method B where these parameters are updated each generation). Due to the small errors and its advantage in solving optimization problems, Method C was used for calculating rates of genetic progress.

Optimization of breeding schemes

An illustration of how the procedure for optimizing breeding schemes works is shown in Table 2. In this example, selection is practised for fourteen generations, $N = 300$, $h^2_{(0)} = 0{\cdot}4$ and $\Delta\bar{G}_{(5,14)}$ is the average rate of gain from generation 5 to generation 14 obtained from (2). Variables N_s and d are optimized by maximizing $\Phi = \Delta\bar{G}_{(5,14)} - \lambda\Delta F$. The parameter λ is varied at intervals of 0·1. If no restrictions on inbreeding are imposed ($\lambda = 0$) then the maximum gain in the last 10 years of selection is obtained when using eight sires and 16 dams. In this situation the rate of inbreeding is around 4%. If for instance, we impose the restriction $\Delta F \leq 2\%$ then the optimum scheme will be that selecting 14 sires and 28 dams ($\lambda = 3{\cdot}0$). Some response is lost over this period (with this restriction on inbreeding, $\Delta\bar{G}_{(5,14)}$ is around 4% less than that obtained with unrestricted inbreeding). The estimate can be refined by varying λ within the interval 2·9 to 3·0.

For restrictions $\Delta F \leq 1\%$, $\Delta F \leq 0{\cdot}5\%$ and $\Delta F \leq 0{\cdot}25\%$, the optimum number of sires is 34, 55 and 85, respectively. Maximum gains under these restrictions are $0{\cdot}417\sigma_{P(0)}$ ($\lambda = 10{\cdot}3$), $0{\cdot}335\sigma_{P(0)}$ ($\lambda = 26{\cdot}3$) and $0{\cdot}234\sigma_{P(0)}$ ($\lambda = 63{\cdot}5$), respectively. As shown in Table 2 the optimum mating ratio is 1 for the most severe restrictions on ΔF.

The coefficient of variation of response resulting from maximizing $\Phi = \Delta\bar{G}_{(5,14)} - \lambda\Delta F$ has been included for illustration in Table 2. However, it should be noted that if it is desired to impose restrictions on $CV_{(5,14)}$ rather than on ΔF, then optimum schemes should be found by following a similar approach when maximizing $\Phi = \Delta\bar{G}_{(5,14)} - \lambda CV_{(5,14)}$.

Optimum schemes for fish breeding with restrictions on ΔF

Comparison of deterministic models with simulation results. This comparison was made for a fixed mating ratio only (i.e. N_s was the only variable optimized). Using stochastic simulation, Gjerde *et al.* (1996) have investigated optimum designs for fish breeding programmes for giving maximum $\Delta\bar{G}_{(5,14)}$ assuming different but fixed values for N, $h^2_{(0)}$ and d. Several restrictions on rates of inbreeding ($\Delta F \leq 0{\cdot}02$,

Table 1 *Comparison between simulated and predicted values for average genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 14 for different heritabilities ($h^2_{(0)}$), mating ratios (d) and number of offspring per mating (n_{off}). In all cases, the numbers of males selected was equal to 20. Values presented are simulated (S) and % errors of predictions P_A, P_B, and P_C (E_{P_A}, E_{P_B}, E_{P_C}) calculated as $E_{P_X} = [(P_X - S)/S] \times 100$*

		$n_{off} = 6$					$n_{off} = 12$				
	$h^2_{(0)}$	0·1	0·2	0·4	0·6	0·9	0·1	0·2	0·4	0·6	0·9
$d = 1$	S	0·0912	0·1751	0·3292	0·4780	0·7048	0·1232	0·2340	0·4366	0·6333	0·9406
	E_{P_A}	12·4	11·5	10·5	8·8	6·5	13·6	13·5	12·8	10·8	7·4
	E_{P_B}	1·6	1·5	1·6	1·1	0·8	1·6	1·8	1·8	1·2	0·5
	E_{P_C}	5·1	3·9	2·6	1·3	0·6	4·8	3·7	2·2	0·8	0·0
$d = 2$	S	0·1116	0·2127	0·3982	0·5752	0·8423	0·1413	0·2669	0·4969	0·7183	1·0630
	E_{P_A}	8·7	8·4	7·6	6·4	4·7	10·1	10·5	9·9	8·2	5·2
	E_{P_B}	0·4	0·6	0·7	0·3	0·2	1·1	1·5	1·5	0·9	0·1
	E_{P_C}	2·9	2·3	1·3	0·4	0·0	3·4	3·0	1·9	0·7	–0·3
$d = 5$	S	0·1319	0·2509	0·4687	0·6752	0·9909	0·1603	0·3040	0·5641	0·8128	1·1968
	E_{P_A}	7·4	7·2	6·4	5·3	3·4	8·2	8·0	7·7	6·3	3·8
	E_{P_B}	0·5	0·6	0·6	0·3	0·2	0·7	0·7	0·9	0·5	–0·2
	E_{P_C}	2·4	1·9	1·0	0·3	–0·4	2·6	1·9	1·2	0·3	–0·5
$d = 10$	S	0·1452	0·2754	0·5145	0·7410	1·0865	0·1722	0·3252	0·6063	0·8741	1·2838
	E_{P_A}	6·7	6·7	5·9	4·8	2·9	7·7	7·9	7·0	5·6	3·3
	E_{P_B}	0·3	0·5	0·4	0·1	–0·4	0·8	1·1	0·7	0·1	–0·4
	E_{P_C}	1·9	1·8	0·8	0·0	–0·6	2·5	2·2	1·0	0·0	–0·6

Table 2 *Average rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 14 ($\Delta\bar{G}_{(5,14)}$), rate of inbreeding (ΔF), coefficient of variation of $\Delta\bar{G}_{(5,14)}$ ($CV_{(5,14)}$), number of sires (N_s) and mating ratio (d) obtained by maximizing the objective function $\Phi = \Delta\bar{G}_{(5,14)} - \lambda\Delta F$ for different values of λ. The number of scored individuals per generation is 300 and the initial heritability is 0·4. Genetic gain is obtained by using P_C*

λ	Φ	$\Delta\bar{G}_{(5,14)}$	ΔF	$CV_{(5,14)}$	N_s	d
0·0	0·498	0·498	0·03912	0·0836	8	2
0·1	0·494	0·497	0·03396	0·0781	9	2
0·2	0·491	0·497	0·03396	0·0781	9	2
0·3	0·487	0·497	0·03396	0·0781	9	2
0·4	0·484	0·497	0·03096	0·0781	9	2
0·5	0·480	0·497	0·03096	0·0781	9	2
0·6	0·477	0·495	0·02988	0·0738	10	2
0·7	0·474	0·495	0·02988	0·0738	10	2
0·8	0·471	0·495	0·02988	0·0738	10	2
2·9	0·419	0·482	0·02158	0·0648	13	2
3·0	0·417	0·476	0·01966	0·0626	14	2
10·2	0·316	0·421	0·01025	0·0517	33	1
10·3	0·315	0·417	0·00983	0·0513	34	1
26·2	0·206	0·338	0·00506	0·0462	54	1
26·3	0·205	0·335	0·00492	0·0461	55	1
63·4	0·079	0·238	0·00251	0·0480	84	1
63·5	0·078	0·234	0·00245	0·0482	85	1

$\Delta F \leq 0{\cdot}01$, $\Delta F \leq 0{\cdot}005$ and $\Delta F \leq 0{\cdot}0025$) were examined. A comparison of their simulation results with predictions obtained with the method described above is shown in Figure 2. In this example, the mating ratio was set to 2. Mean responses over generations 5 to 14 under two restrictions on inbreeding are presented. The value of λ was varied at appropriate intervals for obtaining rates of inbreeding lower than the specified values. Figure 2 shows a good agreement between predictions and simulations although maximum gains obtained with predictions tend to be higher than those obtained with simulations, mostly when N is large. The highest overprediction of gain was around 10%.

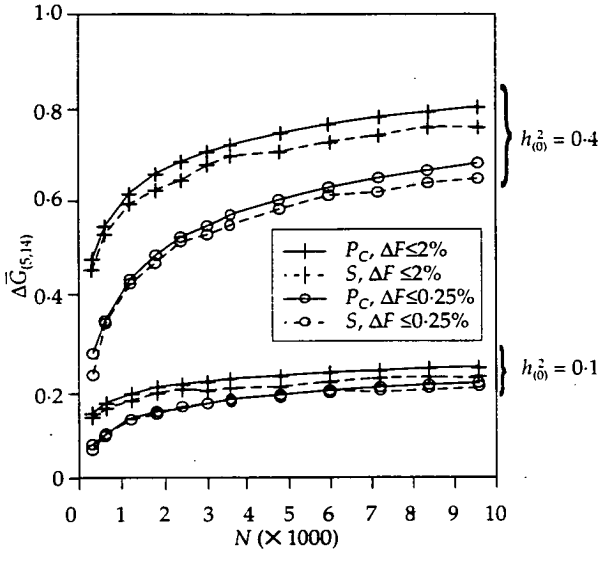


Figure 2 Simulated (S) and predicted (P_C) average rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 14 ($\Delta\bar{G}_{(5,14)}$) achieved with optimum number of sires under two different restrictions on the rate of inbreeding (ΔF) for different scored population sizes (N) and heritabilities ($h^2_{(0)}$). Mating ratio was fixed to 2.

In general, predictions for genetic gain worked better for the lowest values of N and for more stringent restrictions on ΔF, especially when $h^2_{(0)}$ is low (Figure 2). With large numbers of scored individuals and less severe restrictions on inbreeding, selection intensities can become very extreme. In these cases, where the numbers selected become small relative to family sizes, prediction of ΔF is less accurate (Woolliams *et al.*, 1993). Also, selection differentials are likely to be overestimated in the predictions since the correlation of family members has been ignored (Hill, 1976). For $h^2_{(0)} = 0{\cdot}1$, predicted optimum N_s ranged from 12 (N = 300) to 18 (N = 9600) when $\Delta F \leq 2\%$ and from 58 to 111 when $\Delta F \leq 0{\cdot}25\%$. For $h^2_{(0)} = 0{\cdot}4$, optimum N_s ranged from 14 to 25 when $\Delta F \leq 2\%$ and from 60 to 149 when $\Delta F \leq 0{\cdot}25\%$. For the smallest values of N, simulated and predicted optimum N_s were almost equal. However, values for N_s obtained from simulations tend to be higher than those predicted for the largest scored population sizes.

Optimization of schemes for both N_s *and* d. Rates of genetic progress for optimum schemes of different sizes when both N_s and d are optimized are shown in Figure 3. The gain maximized was the average over generations 5 to 14. Different values for $h^2_{(0)}$ and restrictions on inbreeding were considered. Maximum genetic progress obtained with unrestricted inbreeding ($\lambda = 0$) is also shown for comparison. The proportional reduction in genetic gain when imposing constraints on ΔF decreased with N and, at to a lesser extent, with $h^2_{(0)}$. When restricting ΔF to 2%, the loss in response was very small and ranged from 0% ($N > 1800$) to 4% ($N = 300$) for $h^2_{(0)} = 0{\cdot}2$, from 0% ($N > 3000$) to 7% ($N = 300$) for $h^2_{(0)} = 0{\cdot}5$ and from 0% ($N > 3600$) to 7% ($N = 300$) for $h^2_{(0)} = 0{\cdot}8$. With the most severe restriction on ΔF ($\Delta F \leq 0{\cdot}25\%$), the loss in genetic gain ranged from 14% (N =

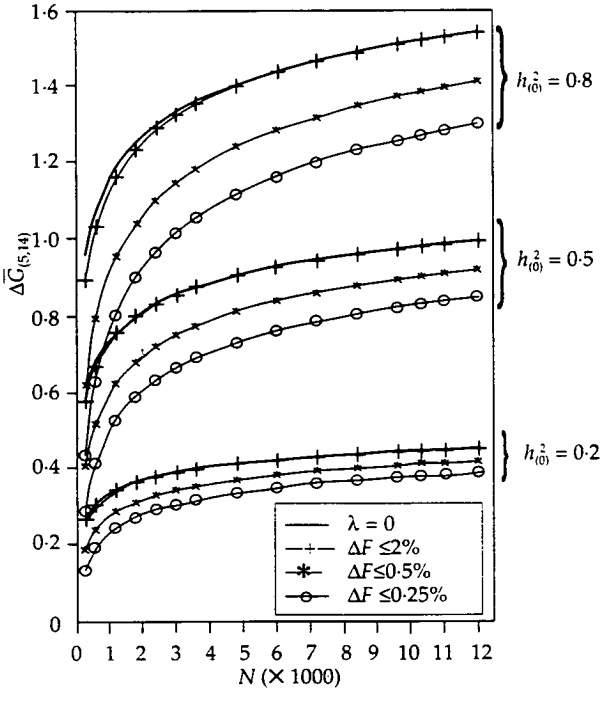


Figure 3 Average rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 14 ($\Delta\bar{G}_{(5,14)}$) achieved with optimum number of sires and mating ratios for different scored population sizes (N), heritabilities ($h^2_{(0)}$) and different restrictions on rate of inbreeding (ΔF). Prediction C (P_C) was used to obtain genetic gain.

12000) to 53% (N = 300) for $h^2_{(0)}$ = 0·2, from 15% (N = 12000) to 53% (N = 300) for $h^2_{(0)}$ = 0·5 and from 16% (N = 12000) to 55% (N = 300) for $h^2_{(0)}$ = 0·8.

For λ = 0 (unrestricted inbreeding), the rate of inbreeding in optimum schemes varied from 2·15% (N = 12000) to 4·32% (N = 300) with $h^2_{(0)}$ = 0·8, from 1·91% to 3·98% with $h^2_{(0)}$ = 0·5 and from 1·79% to 3·43% with $h^2_{(0)}$ = 0·2. For a given N, the proportional reduction in gain by imposing more severe restrictions on ΔF was very similar for different heritabilities. The differences among rates of progress obtained under different constraints of ΔF decrease as N increases. In other words, the gain obtained by increasing N is higher with more severe restrictions on ΔF. More stringent restrictions on ΔF imply lower selection intensities and then more potential for further improvement by increasing N.

Table 3 shows the optimum number of sires and mating ratio for different values of N, heritabilities and restrictions on ΔF. When optimizing for both d and N_s with unrestricted inbreeding, the optimum mating ratio was equal to 2 for all sets of parameters considered. However, with a lower heritability ($h^2_{(0)}$ = 0·1) and small schemes (e.g. N = 300) the optimum mating ratio was d = 1. The optimum number of sires varied from 8 (N = 300) to 24 (N = 12000) for $h^2_{(0)}$ = 0·2, from 8 to 27 for $h^2_{(0)}$ = 0·5 and from 7 to 21 for $h^2_{(0)}$ = 0·8. Larger N_s is required for $h^2_{(0)}$ = 0·5 because, under mass selection, the expected rate of inbreeding (for given N_s and N_d) is maximum at heritabilities close to 0·6. The most severe restriction on ΔF ($\Delta F \leq$ 0·25%) leads to optimum d = 1 when N < 3600 and $h^2_{(0)}$ = 0·2 and when N < 1800 and $h^2_{(0)}$ > 0·2. Under this constraint and for $h^2_{(0)}$ = 0·2, the optimum N_s varied from 82 to 154 for N < 3600 and from 111 to 136 for $N \geq$ 3600. When $h^2_{(0)}$ = 0·8, the optimum N_s varied from 86 to 136 for N < 1800 and from 104 to 143 for $N \geq$ 1800.

Selection intensities were considerably reduced with the most severe restriction on ΔF ($\Delta F \leq$ 0·25%), particularly for the lower values of N. However, they were still very high in most cases. Restricting ΔF to $\Delta F \leq$ 0·25% and for values of N greater than 600, selection intensities were higher than three standard deviation units.

The effect of the mating ratio on genetic progress can be observed in Figure 4 for a scheme of size N = 3000, different heritabilities and restrictions on ΔF. In this case, values are obtained by optimizing N_s with fixed N and d. Genetic progress with unrestricted inbreeding (λ = 0) was very close to that obtained when ΔF restricted to 2%. The proportional decrease in gain by increasing d was very similar for both

Table 3 *Optimum number of sires (N_s) and mating ratio (d) for maximizing average genetic gain between generations 5 and 14 under different constraints on rate of inbreeding (ΔF), heritabilities ($h^2_{(0)}$) and scored population sizes (N). Genetic gain is obtained by using* P_C

		$\lambda = 0$		$\Delta F \leq 2\%$		$\Delta F \leq 0{\cdot}5\%$		$\Delta F \leq 0{\cdot}25\%$	
$h^2_{(0)}$	N	N_s	d	N_s	d	N_s	d	N_s	d
0·2	300	8	2	18	1	52	1	82	1
	600	10	2	15	2	63	1	103	1
	3000	16	2	19	2	61	2	154	1
	6000	20	2	21	2	68	2	122	2
	9600	22	2	22	2	73	2	131	2
	12000	24	2	24	2	75	2	136	2
0·5	300	8	2	15	2	56	1	85	1
	600	11	2	17	2	69	1	111	1
	3000	18	2	22	2	69	2	120	2
	6000	22	2	24	2	78	2	138	2
	9600	25	2	26	2	84	2	151	2
	12000	27	2	27	2	87	2	156	2
0·8	300	7	2	14	2	56	1	86	1
	600	9	2	16	2	68	1	111	1
	3000	15	2	20	2	65	2	115	2
	6000	18	2	21	2	72	2	130	2
	9600	20	2	22	2	76	2	139	2
	12000	21	2	23	2	78	2	143	2

heritabilities. Also, the proportional loss in gain by imposing more severe restrictions on ΔF followed the same pattern for different $h^2_{(0)}$. For schemes of this size the maximum gain was obtained when $d = 2$. For greater values of d, genetic progress decreased with d. By increasing d from 2 to 20, gain was decreased by around 10% with unrestricted inbreeding and by around 20% with the most severe restriction on ΔF.

Optimum schemes for fish breeding with restrictions on $CV_{(5,14)}$

Maximum average rates of genetic progress from generation 5 to generation 14 under different restrictions on variability of response were also analysed. The constraints considered were $CV_{(5,14)} \leq 10\%$ and $CV_{(5,14)} \leq 5\%$. The objective function maximised in this case was $\Phi = \Delta\bar{G}_{(5,14)} - \lambda CV_{(5,14)}$ and the variables optimized were N_s and d. The effects of increasing N and $h^2_{(0)}$ on genetic progress when constrained by $CV_{(5,14)}$ were similar to those described before when the constraints were imposed on ΔF.

For schemes of size $N = 300$, $\lambda = 0$ (i.e. unrestricted $CV_{(5,14)}$) and $h^2_{(0)} = 0{\cdot}8$, $0{\cdot}5$ and $0{\cdot}2$, the coefficient of variation was 4·3%, 7·2% and 11·6%, respectively. Corresponding $CV_{(5,14)}$ for the largest schemes (N = 12000) were 1·8%, 3·0% and 5·1%. Thus, only with the smallest resources and the lowest heritabilities there was some loss in gain when imposing restrictions on $CV_{(5,14)}$. The loss in genetic progress under the restriction $CV_{(5,14)} \leq 10\%$ with respect to unrestricted $CV_{(5,14)}$ was very small in all cases and ranged from 0% ($h^2_{(0)} = 0{\cdot}8$) to 2·2% ($h^2_{(0)} = 0{\cdot}2$ and $N = 300$). Under the restriction $CV_{(5,14)} \leq 5\%$, the relative loss of gain varied from 0% ($h^2_{(0)} = 0{\cdot}8$) to around 30% ($h^2_{(0)} = 0{\cdot}2$ and $N = 600$). Schemes of size $N = 300$ could not achieve the most severe restriction when $h^2_{(0)} = 0{\cdot}2$.

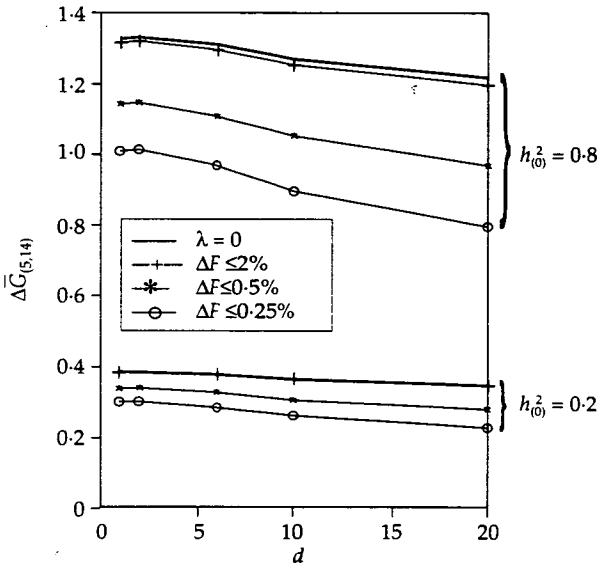


Figure 4 Effect of mating ratio (d) on rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 14 ($\Delta\bar{G}_{(5,14)}$) under different constraints on rate of inbreeding (ΔF) and heritabilities ($h^2_{(0)}$) for a scheme with $N = 3000$. Prediction C (P_C) was used to obtain genetic gain.

Discussion

The results presented have shown how breeding schemes defined only in terms of resources available (N), risk preference (ΔF or CV) and selection criterion ($h^2_{(0)}$) can be designed (in terms of numbers of individuals to be selected) to produce maximum genetic gain. This was achieved by defining simple but appropriate closed expressions for predicting ΔG and ΔF and using a computer routine to maximize a linear function of these in terms of secondary design variables such as the numbers of parents. The results showed that the optimum number of males increased with more severe constraints on ΔF and more resources. In all cases studied the mating ratio was less than or equal to 2 and was equal to 2 when the constraint on ΔF was least severe, when N was large and $h^2_{(0)}$ was higher.

Breeding schemes may be expected to be in place for many generations particularly in species such as fish in which the generation interval can be short. The maximization of short term gains does not necessarily imply maximum long term gains, since the accumulation of inbreeding reduces the potential for future genetic improvement and can reduce performance if the trait under selection is subject to inbreeding depression. Also, inbreeding reduces the fitness of the animal. Furthermore, the variability of response, which determines the predictability of the future genetic mean of the population, can vary in different schemes and the degree of uncertainty may influence decisions on resources. The magnitude of the variance of genetic gain is associated with inbreeding (for example, see the approximation for $V(\Delta G)$ used in this paper).

Two distinct problems arise in the design of breeding programmes with restrictions on risk. The first problem is *a priori*, to maximize response given the basic design variables: available resources, objectives (trait and time scale objectives) and attitude to risk. This will determine the value of additional resources; more or different measures; or the expected benefits and losses from increasing risks. There is a second problem which is *a posteriori* and refers to methods of selection and evaluation for reducing inbreeding in the course of operation of the breeding scheme. The latter methods have received considerable attention in recent years (Toro and Nieto, 1984; Toro and

Perez-Enciso, 1990; Verrier *et al.*, 1993; Grundy *et al.*, 1994; Brisbane and Gibson, 1994; Wray and Goddard, 1994; Villanueva *et al.*, 1994). Various mating designs have also been considered (Woolliams, 1989; Caballero *et al.*, 1996a) and these can be considered to be applicable to both problems. Prediction formulae for different mating designs could be easily incorporated into *a priori* optimization.

The *a posteriori* procedures, proposed in the studies mentioned above, have been efficient in reducing rates of inbreeding with a minimal loss in genetic gain. However, they do not necessarily give maximum gains after a specified number of generations of selection. Wray and Goddard (1994) have used a selection criterion to maximize response corrected for the depression in performance due to inbreeding (D). One advantage of their approach is that the number of animals selected per generation and the number of offspring produced per individual are optimized each generation. This is more difficult to incorporate into *a priori* optimization since predictions are unavailable. Although their objective is to maximize the function $G_t = \sum_{j=1}^{t} \Delta G_j - DF_t$, the selection decisions are made generation by generation. By considering F_t rather than ΔF, the utility function is statically defined by the base generation and becomes less and less relevant over time to the current breeding opportunities. The use of the static base generation and F_t in the objective can alter the rate of inbreeding achieved over time (B. Grundy, unpublished results) and consequently alter the selection rules. The rate of inbreeding is however an intrinsic parameter of a breeding scheme for example for determining the rate of loss of genetic variability (e.g. Meuwissen and Woolliams, 1994c), the fixation probability of mutant genes under selection (Caballero *et al.*, 1996b) and the genetic architecture of inbred chromosomes (Stam, 1980).

In this study, gain is maximized (by optimizing the number of animals selected) when specific restrictions on inbreeding or variability of response are imposed. Although most of the results presented are obtained by maximizing average genetic gain over a specified time period (from generation 5 to generation 14), the method is general for any number of generations of selection. Short- *versus* long-term responses (and optimum schemes in both situations) can be compared easily using this procedure. Also, the method is general with respect to factors to be considered when optimizing breeding schemes. For instance, increasing N leads to greater genetic gains and lower rates of inbreeding. However, the costs of the scheme also increase with N. There can be situations where the extra gain is small and does not compensate for the extra costs (Figure 3). Some function of the cost of the scheme could be incorporated in the objective function to be considered when deciding optimum breeding designs.

The optimization procedure has been applied to a practical situation in fish breeding programmes. The results show that optimum family sizes increased as the resources increased and varied from 10 (N = 300, $h^2_{(0)}$ = 0·5) to 286 (N = 12000, $h^2_{(0)}$ = 0·8) under unrestricted inbreeding. With the rate of inbreeding restricted to 0·25%, optimum family sizes ranged from around 4 (N = 300) to around 40 (N = 12000). Here we have assumed that there is no variation in family size before selection. In fish breeding programmes using mass selection, family sizes are standardized at the eyed egg stage or at the first feeding (when the fish are transferred from the hatchery to feeding tanks). However, there is variation in family size at the time of selection due to variation in survival rate and proportion of sexual mature fish among sib groups (Gjerde *et al.*, 1996). Although this variation has been recognized, it has not been quantified. Variation in family size could be easily included in the predictions and would increase the rate of inbreeding. As a consequence, optimum schemes when restricting ΔF would imply larger numbers of individuals to be selected.

Gjerde *et al.* (1996) have investigated the optimum designs using stochastic simulation and found similar solutions to those described here but their approach is very computationally demanding. For a given set of values for N and $h^2_{(0)}$ and for a given restriction on inbreeding, several simulations must be run to find the optimum solution. The number of selected animals is changed in different simulations until the desired level of inbreeding is achieved. Therefore, designing the scheme *a priori* using computer simulation requires considerable resources, and even then there is no guarantee that the optimum will be found.

The results presented here do not agree with those of Jódar and López-Fanjul (1977) who concluded that the maximum $\Delta G_{(t)}$, for any value of t, is achieved when the selected proportions are the same for both sexes (d = 1). Results from this study with unrestricted inbreeding (λ = 0) indicate optimum mating ratios equal to 2 for the values considered for $h^2_{(0)}$ (0·2, 0·5 and 0·8) and N (300 to 12000). The discrepancy can be ascribed to the prediction of ΔF used in both studies. Jódar and López-Fanjul (1977) predicted ΔF by using $\Delta F = (8N_s)^{-1} + (8N_d)^{-1} = (8N_s)^{-1}(1 + d^{-1})$ which is inappropriate for selected populations. The prediction of ΔF used in this paper is that appropriate for selected populations.

Paper 26

Genet. Res. Camb. (1997), **69**, *pp.* 145–158. *With 4 figures.* Printed in the United Kingdom © 1997 Cambridge University Press

Optimization of breeding programmes under index selection and constrained inbreeding

B. VILLANUEVA[1]* AND J. A. WOOLLIAMS[2]
[1] *Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, Scotland, UK*
[2] *Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland, UK*

(*Received 29 May 1996 and in revised form 3 December 1996*)

Summary

A method for finding optimum breeding schemes which maximize genetic gain under index selection with constraints on the rate of inbreeding is derived. The selection index includes information on the candidate and its sibs. Optimization is for the numbers of males and females to be selected and for the index weights when fixed numbers of offspring per generation, heritabilities and time horizons are considered. The expected rate of gain after a number of generations of selection is combined with the expected asymptotic rate of inbreeding (ΔF) in a single objective function which is maximized for finding the optimum solutions. Under restricted inbreeding, optimum designs are very similar for maximizing gains at different time horizons. The optimum number of selected males (for giving maximum gains) increases with the size of the scheme and with the severity in restricting ΔF and decreases with the heritability. Low heritability, less severe restrictions on ΔF and large schemes lead to increases in the relative weights given to performance of relatives in the index. The presence of common environmental effects leads to increases in optimum mating ratio when the heritability is low, to increases in the number of selected males and to more intense selection within families. Gains from index selection are compared with gains from mass selection. Under restricted inbreeding the advantage of optimized index selection over mass selection is only notable when the heritability is low and the scheme is large (in which case indices put more emphasis on family information than mass selection) and when the heritability is high and the scheme is small (in which case indices put less emphasis on family information).

1. Introduction

In genetic evaluation, information on the performance of relatives of the candidates for selection is used to increase the accuracy of evaluation, and therefore genetic gain. The weights given to records of different relatives in classical selection indices or in BLUP (best linear unbiased prediction) maximize response after a single generation of selection.

The advantage of procedures using information from relatives for improved short-term responses may be offset in the medium or long-term (e.g. Verrier *et al.*, 1993). Firstly, the reduction in genetic variance as a result of linkage disequilibrium generated by selection is higher with more accurate evaluation methods (Bulmer, 1971; Wray & Hill, 1989; Gomez-Raya & Burnside, 1990*a*), although under the infinitesimal model this alone does not dramatically alter the ranking of schemes in the long-term (Wray & Hill, 1989; Woolliams, 1990; Dekkers, 1992; Villanueva & Kennedy, 1993). Secondly, increasing the weights given to relatives' performance in selection decisions increases the rate of inbreeding (Robertson, 1961), particularly when the heritability is low (Belonsky & Kennedy, 1988) which leads to greater reductions in genetic variation. Furthermore a greater rate of inbreeding will reduce the expected fitness of the population and its expected performance in the selected trait if it exhibits inbreeding depression.

Simple mass selection can give higher gains than direct selection on BLUP estimates, which are the most accurate estimates of breeding values using all available information. This has been observed when selection is practised for many generations in small

* Correspondence to be sent to: B. Villanueva, Scottish Agricultural College, Genetics and Behavioural Sciences, Bush Estate, Penicuik EH26 0QE, Scotland UK. Tel: 0131 535 3224; Fax: 0131 535 3121; e-mail: b.villanueva@ed.sac.ac.uk.

populations (Verrier *et al.*, 1993). Also, the advantage of mass selection over BLUP selection after several generations has been shown when the two procedures are compared at the same level of inbreeding (Quinton *et al.*, 1992). The comparisons of within-family and mass selection by Dempfle (1975) showed that long-term response may be greater for within-family selection, particularly when the correlation between the selection criterion and the breeding value is high and selection is intense.

Selection methods that are more sophisticated than within-family and mass selection have been proposed for reducing the emphasis given to family information in the selection criterion and thus reducing the rate of inbreeding with a minimal loss in response (Toro & Perez-Enciso, 1990; Verrier *et al.*, 1993; Grundy *et al.*, 1994; Villanueva *et al.*, 1994; Luo *et al.*, 1995). However, there is no guarantee that the weights given to information from different relatives in these procedures are optimal for maximizing genetic gain after several generations of selection. Optimum index weights for maximizing asymptotic response once the Bulmer equilibrium has been approached were studied by Gomez-Raya & Burnside (1990*b*) by using a selection index which incorporates parental information. The benefits in response from using optimum rather than classical weights were very small. However, they assumed an infinite population and so there was no accumulation of inbreeding.

Wray & Goddard (1994) and Brisbane & Gibson (1995) have proposed methods for selecting individuals aimed at maximizing response when imposing a cost on the inbreeding coefficient. These are *a posteriori* procedures applied once the breeding scheme is in operation and they are effective for offsetting response and inbreeding over a fixed time horizon. Their approach is, however, tied to an identified base population (Villanueva *et al.*, 1996) and the weighting given to family information is not explicitly calculated.

Finding the optimum numbers of individuals to be selected for giving maximum gains is an important problem when designing selection programmes *a priori*. Villanueva *et al.* (1996) have described a procedure for obtaining the optimum numbers of sires and dams for maximizing genetic gain over a specified time period when specific constraints on the rate of inbreeding are imposed. Their method was applied to situations where selection is based on the individual phenotype (i.e. mass selection).

In this paper, this procedure is extended to a situation where selection is on an index which includes information on the individual and its collateral relatives. In addition to the numbers of selected animals, the index weights are optimized by maximizing a single objective function which combined rates of genetic gain and inbreeding. Specific constraints on the rate of inbreeding are considered.

2. Methods

(i) *Model*

The trait under selection is assumed to be determined by an infinite number of additive loci, each with infinitesimal effect (infinitesimal model). A population with discrete generations is assumed with a hierarchical mating structure, in which d dams are mated to each sire and each dam has n offspring. Repeated cycles of directional and truncation selection are practised. Each generation, N_s males and N_d females are selected from $N/2$ males and $N/2$ females scored. The numbers of selected males and females are optimized for each breeding scheme as described later. The number of offspring per mating is determined by N and N_d ($n = N/N_d$). Male and female selection intensities are constant across generations. Mating of selected individuals is at random. A constant number of individuals born per generation (N) is assumed. Different values for N and for the initial heritability ($h^2_{(0)}$) are considered.

Generation 0 consists of unselected individuals with the appropriate family structure. Generation 1 is obtained from the mating of individuals selected at generation 0. The total phenotypic variance at generation t is

$$\sigma^2_{P(t)} = \sigma^2_{A(t)} + \sigma^2_C + \sigma^2_E,$$

where $\sigma^2_{A(t)}$ is the additive genetic variance, σ^2_C is the variance attributed to the common environment of full-sibs and σ^2_E is the individual environmental variance. Environmental variances (σ^2_C and σ^2_E) are constant across generations.

The selection index (I) used as the selection criterion is

$$I = b_1(P - \bar{P}_F) + b_2(\bar{P}_F - \bar{P}_H) + b_3\bar{P}_H,$$

where P is the record of the individual, $\bar{P}_F$ is the mean of n (including the individual) full-sib records, $\bar{P}_H$ is the mean of dn (including the individual and its full-sibs) half-sib records and b_1, b_2 and b_3 are the index weights. This form was used by Wray *et al.* (1994) for predicting the rate of inbreeding under index selection. Note that mass selection is a special case of this index where $b_1 = b_2 = b_3 = 1$. Index weights are assumed constant across generations. Optimum numbers of parents and index weights are obtained by maximizing a single function which combines the expected rates of genetic progress and inbreeding.

(ii) *Prediction of rate of genetic progress*

Prediction of response is obtained (for a given set of N, N_s, N_d, b_1, b_2 and b_3 values) accounting for reduction in genetic variance due to linkage disequilibrium and due to inbreeding. Let $G_{(t)}$ be the average genetic mean of individuals born at generation

t. The rate of genetic gain obtained each generation ($\Delta G_{(t)} = G_{(t)} - G_{(t-1)}$) is

$$\Delta G_{(t)} = i\rho_{(t-1)}\sigma_{A(t-1)},$$

where $i = (0{\cdot}5)\,(i_m + i_f)$ and i_m and i_f are the selection intensities (i.e. standardized selection differentials) for males and females, respectively, and ρ is the accuracy of selection.

The genetic variance is decomposed into between-sire family (σ^2_{As}), between-dam family (σ^2_{Ad}) and within-full-sib family (σ^2_{Aw}) components:

$$\sigma^2_{A(t)} = \sigma^2_{As(t)} + \sigma^2_{Ad(t)} + \sigma^2_{Aw(t)}$$

(Wray & Hill, 1989). These components are obtained each generation by using recurrently the following equations (e.g. Verrier *et al.*, 1990):

$$\sigma^2_{As(t)} = (0{\cdot}25)[1-(1/N_s)][1-k_m\rho^2_{(t-1)}]\,\sigma^2_{A(t-1)},$$
$$\sigma^2_{Ad(t)} = (0{\cdot}25)[1-(1/N_d)][1-k_f\rho^2_{(t-1)}]\,\sigma^2_{A(t-1)},$$
$$\sigma^2_{Aw(t)} = (0{\cdot}5)[1-F_{(t-1)}]\,\sigma^2_{A(0)},$$

where $k_y = i_y(i_y - x_y)$, for $y = m$ (males) of f (females), x_y is the standardized deviation of the truncation point from the mean, $\sigma^2_{A(0)}$ is the value for the genetic variance in the unselected base population and $F_{(t)}$ is the average inbreeding coefficient at generation t. The average coefficient of inbreeding is computed as

$$F_{(t)} = 1-(1-\Delta F)^t,$$

where ΔF is the inbreeding rate. The asymptotic rate of inbreeding (see below) is used to obtain $F_{(t)}$ at each generation.

The accuracy of selection is computed each generation as

$$\rho_{(t)} = \sigma_{AI(t)}/\sigma_{A(t)}\sigma_{I(t)},$$

where $\sigma_{AI(t)}$ is the covariance between the true breeding value and the index and $\sigma_{I(t)}$ is the standard deviation of the index. These are obtained from

$$\sigma_{AI(t)} = b_1\sigma^2_{Aw(t)}\left(1-\frac{1}{n}\right) + b_2\left(\sigma^2_{Ad(t)} + \frac{\sigma^2_{Aw(t)}}{n}\right)\left(1-\frac{1}{d}\right) + b_3\left(\sigma^2_{As(t)} + \frac{\sigma^2_{Ad(t)}}{d} + \frac{\sigma^2_{Aw(t)}}{dn}\right)$$

and

$$\sigma^2_{I(t)} = b_1^2(\sigma^2_{Aw(t)} + \sigma^2_E)\left(1-\frac{1}{n}\right) + b_2^2\left(\sigma^2_{Ad(t)} + \sigma^2_C + \frac{\sigma^2_{Aw(t)} + \sigma^2_E}{n}\right)\left(1-\frac{1}{d}\right) + b_3^2\left(\sigma^2_{As(t)} + \frac{\sigma^2_{Ad(t)} + \sigma^2_C}{d} + \frac{\sigma^2_{Aw(t)} + \sigma^2_E}{dn}\right).$$

(iii) *Prediction of rate of inbreeding*

The asymptotic rate of inbreeding (ΔF) was calculated for a given set of N, N_s, N_d, b_1, b_2 and b_3 values using the expression

$$\begin{aligned}\Delta F = {} & \Delta F_E + \{i^2\alpha_m[(0{\cdot}25)(\tau+\tau_m)^2S^2_\infty - \tau^2_m] \\ & \times[(16N_s)^{-1} + (16N_d)^{-1}]\} \\ & + i^2\alpha_f[(0{\cdot}25)(\tau+\tau_f)^2S^2_\infty - \tau^2_f][8N_d]^{-1} \\ & + Hi^2[2\alpha_m + (1+N_sN_d^{-1})\alpha_f]\{\tau^2B_\infty(S_\infty - 1) \\ & \times[(32N_s)^{-1} + (32N_d)^{-1} - (4N)^{-1}]\} \\ & + Hi^2[2\alpha_m + (1+N_sN_d^{-1})\alpha_f]\{\tau^2S^2_\infty(B_\infty - 1) \\ & \times[(32N_s)^{-1} + (32N_d)^{-1} - (8N)^{-1}]\},\end{aligned}$$

where ΔF_E is the rate of inbreeding assessed assuming independent generations of selection which treats genetic covariances among sibs and between parents and offspring as though they were of environmental origin (i.e. using 'one generation methods'; Wray *et al.*, 1990). In this paper ΔF_E was calculated using eigenvalue methods as described in Appendix A. The other terms are: H, the correction for hypergeometric sampling that is of the form $[1-(0{\cdot}5)N_s^{-1} - (0{\cdot}5)N_d^{-1}]$ (appendix 4 of Woolliams *et al.*, 1993); τ_x, twice the regression of the index of the offspring on the breeding value of the parent of sex x ($\tau_m = b_3$ and $\tau_f = b_2(1-d^{-1}) + b_3d^{-1}$); $\tau = (0{\cdot}5)(\tau_m + \tau_f)$; S_∞ and B_∞ are infinite sums equal to $(1-c)^{-1}$ and $[1-(0{\cdot}5)c^2]$, respectively, where $c = (0{\cdot}5)(1-\tau k\beta)$. In c, $k = (0{\cdot}5)(k_m + k_f)$ and β is the regression of the index on breeding values amongst offspring. Finally, α_x is $(0{\cdot}25)\sigma^2_{Ax(1)}/\sigma^2_{I(2)}$ where $\sigma^2_{Ax(1)}$ is the additive genetic variance after selection in sex x in generation 1.

The method is an index analogy to formula [3] of Woolliams & Thompson (1994) and was used to produce table 1 in that paper. Further validation is given in Appendix B. The accuracy shown is good (up to 8% errors) with an average error of 2·7%.

(iv) *Optimization of breeding schemes*

Optimum schemes are those giving the highest genetic gain for a given rate of inbreeding. The rate of genetic progress at a given generation and the asymptotic rate of inbreeding are obtained for each set of N and $h^2_{(0)}$ values (fixed parameters) as described above. If the objective is to maximize gain over several generations, the average rate of genetic progress from generation $t-1$ to generation $t+m$ ($\Delta\bar{G}_{(t,t+m)}$) is calculated simply as

$$(\Delta G_{(t)} + \Delta G_{(t+1)} + \ldots + \Delta G_{(t+m)})/(m+1).$$

Then optimum schemes are found by maximizing a single objective function (Φ) which combines the rates of genetic gain and inbreeding:

$$\Phi_{(t,t+m)} = \Delta\bar{G}_{(t,t+m)} - \lambda\Delta F.$$

Table 1. *An example of the maximization procedure for* $N = 200$, $h^2_{(0)} = 0{\cdot}3$ *and* $\Phi_{(5,20)} = \Delta\bar{G}_{(5,20)} - \lambda\Delta F$. *Hence, for a restriction of* $\Delta F \leqslant 1\%$, *the scheme for* $\lambda = 7{\cdot}4$ would be expected to give the greatest value of $\Delta\bar{G}_{(5,20)}$ *by using 30 sires* (N_s) *with a mating ratio* (d) *of 1 and a relative weight* ($b_2 = b_3$) *of 1·04 for the family means*

λ	$\Phi_{(5,20)}$	$\Delta\bar{G}_{(5,20)}$	ΔF	N_s	d	$b_2 = b_3$
0·0	0·322	0·322	0·03179	16	1	1·63
1·0	0·295	0·318	0·02336	19	1	1·47
2·0	0·274	0·312	0·01910	21	1	1·33
3·0	0·256	0·304	0·01612	23	1	1·25
7·3	0·201	0·276	0·01020	29	1	1·01
7·4	0·200	0·273	0·00986	30	1	1·04
55·6	−0·009	0·132	0·00253	67	1	0·74
55·7	−0·009	0·130	0·00249	68	1	0·76

This function is denoted as $\Phi_{(t)}$ when $m = 0$ (i.e. when the aim is to maximize the rate of genetic gain at a single generation t). The parameter λ is a Lagrangian multiplier taking positive values and is increased at appropriate intervals until the constraint on ΔF is satisfied. This is then equivalent to maximizing genetic gain with an upper bound on the rate of inbreeding (Woolliams & Thompson, 1994). The procedure has been illustrated in detail by Villanueva *et al.* (1996) for schemes under mass selection.

The index weights (b_1, b_2 and b_3) can be arbitrarily scaled without changing the selection process, so the weight corresponding to the deviation of the individual from the full-sib family mean (b_1) is set to 1. Then, for a given combination of fixed parameters (λ, N and $h^2_{(0)}$) the objective function $\Phi_{(t,t+m)}$ depends on four variables which are N_s, d, b_2 and b_3. The problem is to find the combination of values of these variables which gives the highest value for $\Phi_{(t,t+m)}$. For each possible set of N_s and d values, the optimum index weights were obtained by using the NAG routine E04UCF (The Numerical Algorithms Limited, 1991). When $d = 1$ all sibs are full-sibs and then the selection index is

$$I = b_1(P - \bar{P}_F) + b_2\bar{P}_F \text{ (i.e. } b_2 = b_3).$$

In this situation there are only three variables to be optimized (N_s, d and b_2). For each possible combination of N_s and d values, the optimum index weight was found by using a golden section search in one dimension (e.g. Press *et al.*, 1992).

All possible combinations of N_s and d values (using the optimum weights for each combination) were compared and that set giving the highest value for $\Phi_{(t,t+m)}$ was defined as the optimum. The number of offspring per mating (n) was allowed to be non-integer. Table 1 gives an example of how optimum schemes for maximizing gain under restricted inbreeding were obtained.

Initially, possible values for N_s are between 1 and $N/2$ and possible values for d are between 1 and $N/2N_s$. However, prediction of the rate of inbreeding is inaccurate if selection intensities become very extreme (i.e. with very small numbers of selected individuals). Thus, the minimum number of sires was set to 10, 20 and 30 for N equal to 200, 800 and 3200, respectively (i.e. smaller scored population sizes allowed to have a lower minimum N_s for accurate prediction).

(v) *Selection limits with the infinitesimal model*

The optimization procedure can be used to find optimum selection proportions and index weights for maximizing ultimate response at the selection limit. Robertson (1960), Jódar & López-Fanjul (1977) and Cockerham & Burrows (1980) found that the optimum selection proportion for obtaining the maximum advance at the limit is the same in both sexes ($d = 1$) and equals 1/2. These expectations ignored the Bulmer effect and assumed inbreeding rates only appropriate for populations under random selection. Under mass selection, when using ΔF appropriate for selected populations and accounting for the Bulmer effect, the optimal proportions are somewhat higher (Woolliams & Pong-Wong, 1995).

Optimum schemes (selected proportions and index weights) for maximizing ultimate response under index selection were obtained as described in the previous section but using the objective function $G_{(t)} = \Delta G_{(1)} + \ldots + \Delta G_{(t)}$ and choosing a value of t such that $G_{(t)} - G_{(t-1)}$ is less than 0.01.

An explicit expression for the selection plateau ($G_{(\infty)}$) can be obtained when selection is based only on the individual's own measurement. Under mass selection, the rate of genetic response at any generation t can be approximated as

$$\Delta G_{(t)} \approx \frac{h^2_{(0)} i(a^{t+1} - b^{t+1})}{2\sigma_{P(L)} a^3(a-b)},$$

where $b = (0{\cdot}5)(1 - kh^2_{(L)})$, $a = (1-\Delta F)/\{1 - [0{\cdot}25)h^2_{(0)}\Delta F]/\sigma^2_{P(L)}\}$ and $h^2_{(L)}$ and $\sigma_{P(L)}$ represent Bulmer equilibrium values for the heritability and phenotypic standard deviation (Villanueva *et al.*, 1996). Cumulative gain at the selection limit ($t \to \infty$) is then

$$G_{(\infty)} = \sum_{t=1}^{\infty} \Delta G_{(t)} \approx \frac{h^2_{(0)} i[a^2(1-a)^{-1} - b^2(1-b)^{-1}]}{2\sigma_{P(L)} a^3(a-b)}.$$

(vi) *Computer simulation*

Stochastic simulation was used to test some of the results obtained. In general, the simulation procedure used was that described in Villanueva *et al.* (1996) for mass selection with some modifications: (1) With index selection an extra generation (generation 00) needed to be generated to create the base generation

with family structure (generation 0). Generation 00 was constituted by N_s males and N_d females and these unrelated individuals were mated at random to create generation 0. (2) The phenotypic value of an individual was generated as the sum of its genetic value, an environmental component common to its full-sibs and an individual environmental component. The common and individual environmental components were obtained from normal distributions with mean zero and variance σ_c^2 and $1-h_{(0)}^2-\sigma_C^2$, respectively. Environmental variances were maintained constant over generations. (3) The selection criterion was the selection index described above. (4) Five thousand replicates were run for each simulation.

3. Results

(i) *Optimum schemes under index selection with* $\sigma_c^2 = 0$

(a) *Number of individuals to be selected*

Table 2 shows the optimum numbers of males to be selected and optimum mating ratios for obtaining maximum rate of genetic gain at generation 5 ($\Delta G_{(5)}$) or 20 ($\Delta G_{(20)}$) under two constraints on the rate of inbreeding ($\Delta F \leqslant 1\%$ and $\Delta F \leqslant 0{\cdot}25\%$) and for obtaining maximum $\Delta G_{(20)}$ with unrestricted inbreeding ($\lambda = 0$). The optimum number of selected males increased with the size of the scheme and with more severe constraints on ΔF and decreased with the heritability (over the range of $h_{(0)}^2$ considered). The optimum mating ratio was equal to 1 except for the larger schemes with less severe restrictions on ΔF and greater heritabilities. In these cases the optimum d was 2. The optimum number of offspring per mating ranged from around 3 ($N = 200$, $h_{(0)}^2 = 0{\cdot}1$, $\Delta F \leqslant 0{\cdot}25\%$) to around 43 ($N = 3200$, $h_{(0)}^2 = 0{\cdot}6$, $\lambda = 0$).

With unrestricted inbreeding ($\lambda = 0$) the optimum N_s increased substantially when maximizing $\Delta G_{(20)}$ compared with maximizing $\Delta G_{(5)}$. For $t = 5$, the optimum N_s was constrained by the imposed lower bound on N_s (see Methods) and results are not shown in Table 2. With ΔF restricted, the optimum N_s for maximizing gain was similar for the two time points chosen, with identical mating ratios. Maximization of the total gain from generation 5 to 20 was also examined, and optimum N_s was always between the optimum numbers for maximizing $\Delta G_{(5)}$ and $\Delta G_{(20)}$ (results not shown).

Male selection intensity (i_m) increased with N and $h_{(0)}^2$ and decreased with the severity of the restriction on ΔF and ranged from 0·51 ($N = 200$, $h_{(0)}^2 = 0{\cdot}1$, $\Delta F \leqslant 0{\cdot}25\%$) to 2·37 ($N = 3200$, $h_{(0)}^2 = 0{\cdot}6$, $\lambda = 0$). Female selection intensities (i_f) followed the same trends although they did not always increase with $h_{(0)}^2$ and N because of shifts in d. Selection intensity in females ranged from 0·51 ($N = 200$, $h_{(0)}^2 = 0{\cdot}1$, $\Delta F \leqslant 0{\cdot}25\%$) to 2·09 ($N = 3200$, $h_{(0)}^2 = 0{\cdot}6$, $\lambda = 0$).

(b) *Index weights*

The relative weights given to sib information decreased with heritability over the range considered and with more severe restrictions on ΔF and increased with the size of the scheme (Table 3). The weights given to family information when maximizing $\Delta G_{(20)}$ under unrestricted inbreeding ($\lambda = 0$) were considerably lower than the classical index weights which maximize one-generation gain ($t = 1$) for the same values of N_s and d.

For high $h_{(0)}^2$ and large N and for the time points chosen, the optimum weights were close to those corresponding to mass selection ($b_1 = b_2 = b_3 = 1$). For high $h_{(0)}^2$ and small N, the optimum weights move further towards within-family selection and this is potentiated by restricting ΔF.

In general, the index weights corresponding to sib information were slightly higher when maximizing

Table 2. *Optimum number of sires* (N_s) *and mating ratios* (d) *under index selection for maximizing genetic gain at generations* $t = 5$ *and* $t = 20$ *under different constraints on the rate of inbreeding* (ΔF), *heritabilities* ($h_{(0)}^2$) *and scored population sizes* (N)

		$\lambda = 0$		$\Delta F \leqslant 1\%$				$\Delta F \leqslant 0{\cdot}25\%$			
		$t = 20$		$t = 5$		$t = 20$		$t = 5$		$t = 20$	
$h_{(0)}^2$	N	N_s	d	N_s	d	N_s	d	N_s	d	N_s	d
0·1	200	22	1	32	1	32	1	69	1	69	1
	800	46	1	59	1	61	1	126	1	126	1
	3200	91	1	107	1	111	1	233	1	235	1
0·3	200	21	1	29	1	30	1	68	1	68	1
	800	29	2	35	2	37	2	116	1	117	1
	3200	53	2	58	2	61	2	140	2	142	2
0·6	200	17	1	24	1	25	1	65	1	65	1
	800	22	2	27	2	29	2	96	1	98	1
	3200	37	2	38	2	41	2	107	2	109	2

Table 3. *Optimum index weights (b_2 and b_3) when maximizing genetic gain at generations $t = 5$ and $t = 20$ under different constraints on the rate of inbreeding (ΔF) and index weights obtained under standard selection index theory ($t = 1$) for schemes of different sizes (N) and heritabilities ($h^2_{(0)}$). Index weight b_1 is equal to 1 in all cases. Index weight b_3 is shown (in brackets) only in the cases where it is different from b_2 (i.e. when $d \neq 1$)*

		$\lambda = 0$				$\Delta F \leqslant 1\%$				$\Delta F \leqslant 0{\cdot}25\%$			
		$t = 1$		$t = 20$		$t = 5$		$t = 20$		$t = 5$		$t = 20$	
$h^2_{(0)}$	N	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)	b_2	(b_3)
0·1	200	6·83		2·12		1·53		1·60		1·07		1·07	
	800	9·60		2·97		2·29		2·43		1·50		1·52	
	3200	12·69		3·90		3·26		3·43		2·24		2·28	
0·3	200	3·93		1·43		0·97		1·06		0·75		0·76	
	800	3·56	(4·66)	1·87	(1·82)	1·52	(1·37)	1·63	(1·54)	0·95		0·98	
	3200	4·39	(5·15)	2·26	(2·09)	2·00	(1·80)	2·14	(1·98)	1·50	(1·34)	1·53	(1·38)
0·6	200	2·11		0·78		0·48		0·54		0·45		0·45	
	800	2·06	(2·23)	1·17	(0·90)	0·97	(0·67)	1·06	(0·77)	0·47		0·49	
	3200	2·20	(2·29)	1·26	(0·97)	1·11	(0·81)	1·22	(0·93)	0·93	(0·63)	0·97	(0·66)

Table 4. *Rates of inbreeding × 100 (ΔF) and genetic gain ($\sigma_{P(0)}$ units) at generation 20 ($\Delta G_{(20)}$) when maximizing $\Delta G_{(20)}$ under unrestricted inbreeding ($\lambda = 0$) and rates of genetic gain at generations 5 ($\Delta G_{(5)}$), 20 and average gain between generations 5 and 20 ($\Delta \bar{G}_{(5,20)}$) when maximizing respectively $\Delta G_{(5)}$, $\Delta G_{(20)}$ and $\Delta \bar{G}_{(5,20)}$ under two constraints on ΔF, different heritabilities ($h^2_{(0)}$) and scored population sizes (N)*

		$\lambda = 0$		$\Delta F \leqslant 1\%$			$\Delta F \leqslant 0{\cdot}25\%$		
$h^2_{(0)}$	N	ΔF	$\Delta G_{(20)}$	$\Delta G_{(5)}$	$\Delta G_{(20)}$	$\Delta \bar{G}_{(5,20)}$	$\Delta G_{(5)}$	$\Delta G_{(20)}$	$\Delta \bar{G}_{(5,20)}$
0·1	200	2·09	0·109	0·113	0·100	0·107	0·049	0·047	0·048
	800	1·73	0·166	0·181	0·159	0·169	0·118	0·114	0·116
	3200	1·48	0·226	0·250	0·223	0·236	0·185	0·179	0·182
0·3	200	2·00	0·278	0·292	0·258	0·274	0·132	0·128	0·130
	800	1·54	0·393	0·430	0·385	0·407	0·299	0·290	0·294
	3200	1·25	0·507	0·560	0·504	0·532	0·439	0·426	0·433
0·6	200	1·96	0·525	0·546	0·490	0·517	0·252	0·245	0·248
	800	1·48	0·715	0·779	0·702	0·741	0·558	0·542	0·549
	3200	1·16	0·894	0·995	0·892	0·939	0·794	0·772	0·781

$\Delta G_{(20)}$ than when maximizing $\Delta G_{(5)}$ under specific restrictions on ΔF. If the numbers of selected individuals were fixed, the contrary would be expected (long-term response would be more affected by accumulation of inbreeding and therefore less weight would be given to family information). However, the numbers of individuals to be selected were optimized here and they differ when maximizing gain at different generations. Selection intensities were slightly higher when maximizing early responses under restricted ΔF (Table 2).

(c) *Rates of genetic gain and inbreeding*

Genetic gain at generations 5 and 20 and average response between generations 5 and 20 under different restrictions on ΔF are shown in Table 4. Genetic progress and rate of inbreeding when maximizing $\Delta G_{(20)}$ with unrestricted inbreeding ($\lambda = 0$) are also presented. With unrestricted inbreeding, ΔF was relatively constant for a given value of N. As expected, the rate of genetic progress increased as N and $h^2_{(0)}$ increased and decreased when restrictions on ΔF were imposed. When maximizing $\Delta G_{(20)}$ the proportional reduction in response below the maximum possible ($\lambda = 0$) was small with the least severe restriction on ΔF ($\Delta F \leqslant 1\%$), ranging from 0·2% ($N = 3200$, $h^2_{(0)} = 0{\cdot}6$) to 8·3% ($N = 200$, $h^2_{(0)} = 0{\cdot}1$). With the most severe restriction ($\Delta F \leqslant 0{\cdot}25\%$), gain was from 13·4% to 53·0% lower than that obtained for $\Delta F \leqslant 1\%$. The loss in gain from restricting inbreeding was greater when short-term response ($t = 5$) was maximized judged by the comparisons of response for $\Delta F \leqslant 1\%$ and $\Delta F \leqslant 0{\cdot}25\%$.

Higher rates of gain were obtained when maximizing $\Delta G_{(5)}$ than when maximizing $\Delta G_{(20)}$ due to the higher

Table 5. *Average rate of genetic gain ($\sigma_{P(0)}$ units) between generations 5 and 20 ($\Delta\bar{G}_{(5,20)}$) achieved with optimum number of sires (N_s), mating ratios (d) and index weights (b_2 and b_3) when the rate of inbreeding is constrained to 1% for different heritabilities ($h^2_{(0)}$), scored population sizes (N) and common environment variances (σ^2_C). Index weight b_1 is equal to 1 in all cases. Index weight b_3 is shown (in brackets) only in the cases where it is different from b_2 (i.e. when $d \neq 1$)*

		$\sigma^2_C = 0{\cdot}05$					$\sigma^2_C = 0{\cdot}20$				
$h^2_{(0)}$	N	$\Delta G_{(5,20)}$	N_s	d	b_2	(b_3)	$\Delta G_{(5,20)}$	N_s	d	b_2	(b_3)
0·1	200	0·103	31	1	1·33		0·096	28	1	0·80	
	800	0·158	40	2	1·58	(2·09)	0·144	30	3	0·87	(1·42)
	3200	0·212	59	3	1·95	(2·98)	0·189	40	6	0·96	(2·16)
0·6	200	0·524	23	1	0·38		0·572	20	1	0·17	
	800	0·744	32	1	0·41		0·815	24	1	0·14	
	3200	0·935	32	2	0·62	(0·53)	1·019	30	1	0·18	

accumulation of inbreeding in the latter situation. The maximum average gain over generations 5 and 20 was always between the maximum $\Delta G_{(5)}$ and $\Delta G_{(20)}$. The value of λ for achieving maximum gains at $t = 20$ under specific constraints on ΔF ranged from 1·7 ($\Delta F \leqslant 1\%$, $h^2_{(0)} = 0{\cdot}1$, $N = 3200$) to 107·0 ($\Delta F \leqslant 0{\cdot}25\%$, $h^2_{(0)} = 0{\cdot}6$, $N = 200$).

(ii) *Optimum schemes under index selection with $\sigma^2_C \neq 0$*

The effect of including a non-zero common environmental variance in the model on the optimum numbers of individuals to be selected and index weights and on the maximum rates of genetic gain under restricted inbreeding is shown in Table 5. Common environmental effects led to more emphasis on selection within families, more intense selection on sires and, for low heritability and large schemes, less intense selection on females (see also Table 2 for $\sigma^2_C = 0$). Male selection intensities when restricting ΔF to 1% ranged from 1·20 ($N = 200$, $h^2_{(0)} = 0{\cdot}1$) to 2·44 ($N = 3200$, $h^2_{(0)} = 0{\cdot}6$) when $\sigma^2_C = 0{\cdot}2$ compared, respectively, with 1·12 to 2·36 when $\sigma^2_C = 0{\cdot}0$.

The optimum mating ratio increased when $\sigma^2_C \neq 0$. As in the situation where $\sigma^2_C = 0$, the highest mating ratios were obtained with the least severe restrictions on ΔF (results not shown) and with the largest scored population sizes. However, with common environmental effects the highest d occurred with the lowest heritabilities. With $h^2_{(0)} = 0{\cdot}1$ and $N = 3200$, the optimum d was as large as 6. Fig. 1 shows the rate of gain at generation 20 with unrestricted inbreeding for different mating ratios and $\sigma^2_C = 0{\cdot}0$ or 0·2. In this case, d is fixed and the variables optimized are N_s, b_2 and b_3. Although differences in gain with different mating ratios are small, the optimum d is clearly different from 1 or 2 when $\sigma^2_C = 0{\cdot}2$.

A further check that $d = 1$ may be a long way from the optimum was made by simulation using two

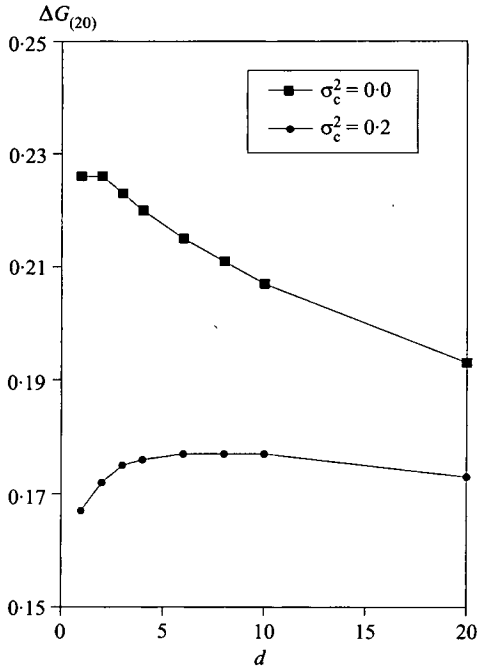


Fig. 1. Effect of mating ratio (d) on rate of genetic gain at generation 20 ($\Delta G_{(20)}$) under unrestricted inbreeding for different values of the variance of common environmental effects (σ^2_C). The total number of individuals scored is 3200, the heritability is 0·1 and the number of selected males and the index weights are optimized for maximizing $\Delta G_{(20)}$ for fixed d.

schemes with equal and integer numbers of offspring per family, $N = 3200$, $h^2_{(0)} = 0{\cdot}1$, $\sigma^2_C = 0{\cdot}20$ and $\lambda = 0$ that were compared for $\Delta G_{(20)}$. Scheme A was close to the optimum (as determined by the deterministic model) for fixed $d = 1$ ($N_s = 80$ and $b_2 = b_3 = 1{\cdot}07$) and scheme B was close to the optimum for unrestricted d ($N_s = 40$, $d = 8$, $b_2 = 1{\cdot}05$ and $b_3 = 2{\cdot}79$). The simulated $\Delta G_{(20)}$ were 0·164 and 0·174 for schemes A and B, respectively, which are remarkably close to those predicted (0·167 and 0·176).

Table 6. *Comparison of index selection with mass selection relative to male (D_{im}) and female (D_{if}) selection intensities when maximizing genetic gain at generation 20 under different constraints on the rate of inbreeding (ΔF), heritabilities ($h^2_{(0)}$) and scored population sizes (N). Values presented are differences of results from index selection minus results from mass selection expressed as a percentage of results from mass selection*

		$\lambda = 0$		$\Delta F \leqslant 1\%$		$\Delta F \leqslant 0{\cdot}25\%$	
$h^2_{(0)}$	N	D_{im}	D_{if}	D_{im}	D_{if}	D_{im}	D_{if}
0·1	200	−7·5	−7·5	−8·2	−8·2	0·0	0·0
	800	−9·6	−9·6	−14·8	−14·8	−8·9	−8·9
	3200	−11·5	−11·5	−15·4	−15·4	−13·7	−13·7
0·3	200	−4·2	−4·2	0·0	0·0	6·0	6·0
	800	−4·0	−6·0	−5·8	−8·3	0·9	0·9
	3200	−5·1	−6·3	−7·7	−9·2	−3·2	−5·2
0·6	200	2·1	2·1	11·4	11·4	21·3	21·3
	800	0·0	0·0	2·2	3·3	12·3	12·3
	3200	0·0	0·0	0·4	0·5	3·8	4·6

Table 7. *Comparison of index selection with mass selection at generation 20 relative to rates of inbreeding ($D_{\Delta F}$) and genetic gain ($D_{\Delta G_{20}}$) when inbreeding is unrestricted ($\lambda = 0$) and at generations 5 and 20 relative to rates of genetic gain ($D_{\Delta G_5}$, $D_{\Delta G_{20}}$) when the rate of inbreeding is restricted to 1% or 0·25% for different heritabilities ($h^2_{(0)}$) and scored population sizes (N). Values presented are differences of results from index selection minus results from mass selection expressed as a percentage of results from mass selection*

		$\lambda = 0$		$\Delta F \leqslant 1\%$		$\Delta F \leqslant 0{\cdot}25\%$	
$h^2_{(0)}$	N	$D_{\Delta F}$	$D_{\Delta G_{20}}$	$D_{\Delta G_5}$	$D_{\Delta G_{20}}$	$D_{\Delta G_5}$	$D_{\Delta G_{20}}$
0·1	200	31	8	3	4	2	0
	800	54	16	11	12	3	3
	3200	72	25	21	23	10	10
0·3	200	10	1	1	0	0	0
	800	19	3	1	2	0	0
	3200	26	5	4	5	1	1
0·6	200	−6	0	3	2	6	6
	800	−4	0	1	0	3	3
	3200	1	0	1	0	1	1

The emphasis given to family means was notably reduced with the largest σ^2_C ($\sigma^2_C = 0{\cdot}2$) in comparison with equivalent results for $\sigma^2_C = 0{\cdot}0$ (see Table 3 for $\sigma^2_C = 0{\cdot}0$), leading to more emphasis upon selection within families, particularly with high $h^2_{(0)}$. Also, when optimum d is different from 1 and σ^2_C is large ($\sigma^2_C = 0{\cdot}2$), the weight given to the half-sib family mean was always higher than the weight given to the full-sib mean, putting more emphasis upon the sire information. The rate of gain was in general decreased when including common environmental effects, except for high $h^2_{(0)}$ (see Table 4 for $\sigma^2_C = 0{\cdot}0$), where gain was greatest with the highest σ^2_C considered ($\sigma^2_C = 0{\cdot}2$).

(iii) *Mass selection compared with index selection*

(a) *Numbers of individuals to be selected*

A comparison of optimum schemes for mass selection, modelled by setting $b_1 = b_2 = b_3 = 1$, and optimum schemes with indices is shown in Table 6. With unrestricted ΔF, the greatest differences in male selection intensity between the two selection methods were with the lowest heritability and the large scored population sizes. There were no general rules concerning selection intensity: it was higher with mass than with index selection when $h^2_{(0)} = 0{\cdot}1$, but lower with mass selection when $h^2_{(0)} = 0{\cdot}6$.

(b) *Rate of inbreeding when maximizing $\Delta G_{(20)}$ for $\lambda = 0$*

At low heritabilities, ΔF is substantially higher with index than with mass selection (Table 7). However, at $h^2_{(0)} = 0{\cdot}6$ mass selection gave greater ΔF for the smallest schemes, despite the fact the optimum selection intensities were lower (Table 6). The higher ΔF obtained with mass selection is due to the fact that under these circumstances the optimum index approaches within-family selection (see Table 3).

Figure 2 shows the effect of the heritability on rates of inbreeding obtained for optimum schemes under mass and index selection. For maximizing gain at generation 20, the optimum schemes under index selection maintained a relatively constant ΔF (around 0·02 for $N = 200$ and 0·012 for $N = 3200$) for all but extreme heritabilities (see also Table 4). As $h^2_{(0)}$ tends to 0 or 1, ΔF was increased. In contrast, for mass selection the optimum schemes for maximizing $\Delta G_{(20)}$ showed ΔF increased from 0·015 to 0·028 for $N = 200$ and from 0·007 to 0·015 for $N = 3200$ as $h^2_{(0)}$ moved from 0 to 1. The crossover points with index selection were around 0·5 for $N = 200$ or 0·6 for $N = 3200$, at which the optimum index schemes were close to mass selection.

Optimizing the scheme design together with the weights (as described above) was contrasted to optimizing only the weights for a specific N_s ($N_s = 20$) and d ($d = 1$). Fig. 3 shows the rates of inbreeding obtained when using optimum weights for maximizing gains at generations, 1, 5 or 20 without a constraint on inbreeding. The curves for I_{20} and M may be directly

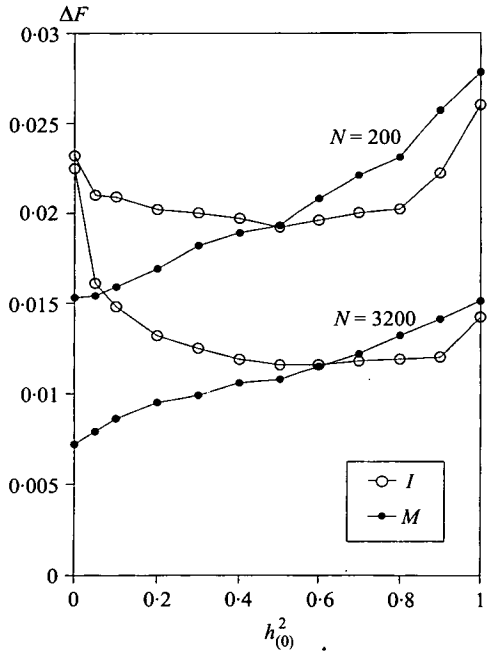


Fig. 2. Asymptotic rate of inbreeding (ΔF) using optimum schemes for maximizing genetic gain at generation 20 under mass (M) and index selection (I) for different scored population sizes (N) and heritabilities ($h^2_{(0)}$).

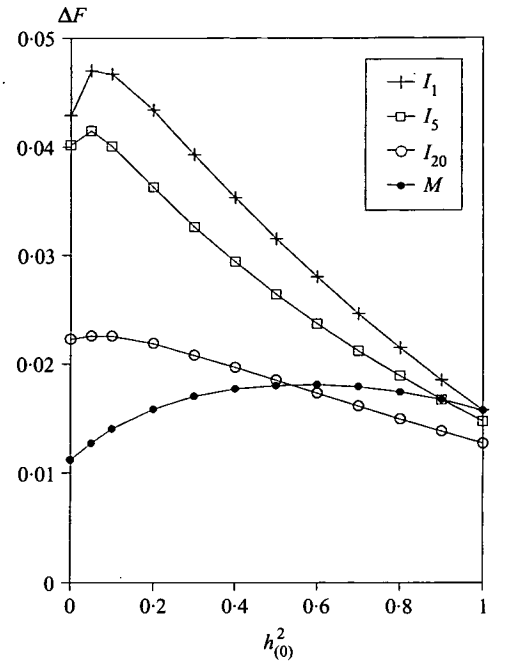


Fig. 3. Asymptotic rate of inbreeding (ΔF) under mass (M) and index selection using optimum weights for maximizing genetic gain at generation 1 (I_1), 5 (I_5) or 20 (I_{20}) for fixed N_s and d ($N_s = 20$, $d = 1$) and different heritabilities ($h^2_{(0)}$). The total number of individuals scored is 200.

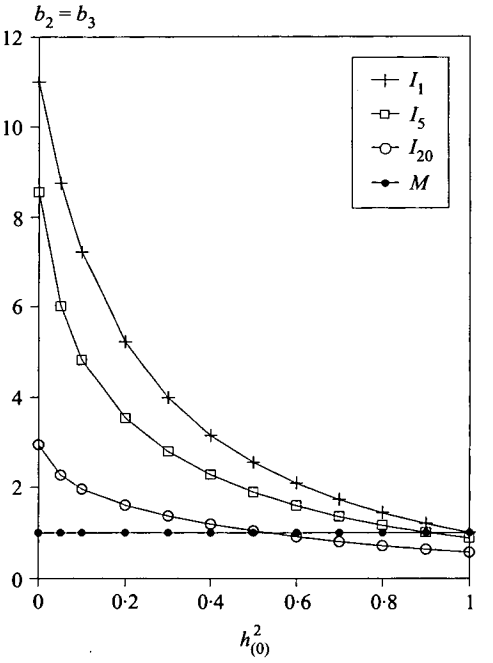


Fig. 4. Optimum weights ($b_2 = b_3$) under mass (M) and index selection for maximizing genetic gain at generations 1 (I_1), 5 (I_5) or 20 (I_{20}) for fixed N_s and d ($N_s = 20$, $d = 1$) and different heritabilities ($h^2_{(0)}$). A weight of 1 is assigned to within-full-sib family deviations ($b_1 = 1$). The total number of individuals scores is 200.

compared with Fig. 2. Although the curve for I_{20} is more constant than those for I_1 and I_5 it still retains the classical shape for indices characterized by the curve for I_1. Likewise the curve for M is quite different from Fig. 2. Fig. 4 shows for $N_s = 20$ and $d = 1$ the

reduction in the weights given to family information as the time point of interest increases, compared with the classical weights maximizing gain in the next generation.

(c) *Rate of genetic progress*

Table 7 shows for the cases considered that unless the heritability is low ($h^2_{(0)} = 0{\cdot}1$) and the scheme is large, the benefits from index selection are small. The gains from indices do not always increase with N. For example, when $h^2_{(0)} = 0{\cdot}6$, the benefit from indices was larger with $N = 200$. This pattern is related to the pattern of optimum weights in Table 3. Since mass selection is an intermediate point between the classical index selection ($b_2 > 1$ and $b_3 > 1$) and within-family selection ($b_2 = b_3 = 0$) it should be expected that the benefits of index selection will be related to what degree the optimum index weights deviate from $b_2 = b_3 = 1$.

When ΔF is unrestricted and conventional weights are used, the advantage of index over mass selection decreases with the number of generations of selection due to the higher accumulation of inbreeding. However, results from Table 7 show that when ΔF is restricted and the weights are optimized, the advantage of index selection can be greater when maximizing long-term responses ($\Delta G_{(20)}$) than when maximizing early responses ($\Delta G_{(5)}$).

(iv) *Selection limits with the infinitesimal model*

The optimum proportions of selected males (p_m) and females (p_f) for maximizing the selection plateau were around 2/3 for all the different values of $h^2_{(0)}$ and N studied. The optimum d was 1 and so $p_m = p_f = p$ and $b_2 = b_3$. For a given heritability, the optimum index weights were approximately constant for different values of N. Optimum b_2 and b_3 were 1·2 for $h^2_{(0)} = 0{\cdot}1$, 1·0 for $h^2_{(0)} = 0{\cdot}3$ and 0·7 for $h^2_{(0)} = 0{\cdot}6$. The theoretical maximum response ranged from 20·9 ($h^2_{(0)} = 0{\cdot}1$ and $N = 200$) to 2023·1 ($h^2_{(0)} = 0{\cdot}6$ and $N = 3200$). The number of generations required to approach the limit increased with N. The time scale of the response is proportional to the effective population size (Robertson, 1960), which was larger for large N (Table 4).

Computer simulations were used to check that b_2 and b_3 clearly differ from zero (i.e. optimum selection differs from within-family selection) for maximizing the selection limit. For fixed $N = 100$, $h^2_{(0)} = 0{\cdot}1$, $N_s = 25$ and $d = 1$, the optimum weights are $b_2 = b_3 = 0{\cdot}85$ and the predicted $G_{(\infty)}$ (from the analytical procedure) is 9·59. For the same set of parameters but $b_2 = b_3 = 0$, the predicted $G_{(\infty)}$ is 6·12. Four thousand generations were run for both schemes and the simulated $G_{(\infty)}$ were 9·80 (standard error = 0·01) and 6·03 (standard error = 0·01) for $b_2 = b_3 = 0{\cdot}85$ and $b_2 = b_3 = 0$, respectively.

4. Discussion

The procedure described here optimizes both the numbers selected and the relative index weights for maximizing responses at a set of time points with or without constraints upon the rate of inbreeding. This is distinct from much of the previously published work comparing selection upon indices and phenotypes in which fixed numbers of individuals were selected and where classical weights were given to family information (Belonsky & Kennedy, 1988; Verrier *et al.*, 1993; Wei *et al.*, 1996). Comparison of different methods at similar rates of inbreeding has been an approach adopted previously by, for example, Woolliams (1989) and Wray & Simm (1990), and was used here since ΔF is an intrinsic genetic property of a scheme determining rate of loss of genetic variation, fixation probabilities of mutants under selection (Caballero *et al.*, 1997) and the genetic architecture of inbred chromosomes (Stam, 1980). The constraint on ΔF can be viewed as the measure of genetic risk, and in this context critical values have been advanced by Meuwissen & Woolliams (1994) amongst others.

The study coming closest to the situation considered here was that of Quinton *et al.* (1992), which was based upon stochastic computer simulation. They allowed a variable number of sires and found, for low levels of inbreeding, higher selection responses from mass selection than from selection based upon BLUP when the two procedures were compared over 20 generations at the same rate of inbreeding. Their results from a selection index including the individual, full-sibs and half-sibs were very similar to those obtained from BLUP, and so their results are comparable with those presented here. For a trait with $h^2_{(0)} = 0{\cdot}25$ (see their table 2), $N = 200$, and a fixed number of dams ($N_d = 50$), with the rate of inbreeding $\approx 1\,\%$ (cumulative inbreeding = 0·18) the cumulative response was higher for mass selection (4·43 σ_P units) than for index selection (3·73 σ_P units) and the optimum number of sires was higher when selection was based upon BLUP than with mass selection (36 versus 18). Our results show that under these specific conditions mass selection was very close to the optimum since the weights under optimized index selection are $b_1 = 1$ and $b_2 = b_3 = 1{\cdot}06$, and ΔG and N_s are practically the same for mass and optimized index selection (Tables 6, 7).

BLUP selection can be closely approximated by using a selection index including the estimated breeding values of the sire and the dam and the mean of estimated breeding values of all dams mated to the sire in addition to information on the individual and its sibs (Wray & Hill, 1989). The procedure used here could be extended to BLUP selection by optimizing three extra index weights. Predictions of the rate of inbreeding are needed, however. Optimization becomes more complex as more sources of information are included in the index. However, results from

BLUP selection would be expected to show very similar trends to those presented here, which was the conclusion of Quinton *et al.* (1992).

When optimizing not only numbers selected but also the weights given to family information, methods using information on relatives must be always equal or superior to mass selection. The general framework considered here makes it clear that the adequacy of mass selection for giving gains close to the maximum will depend entirely upon the time horizon, total offspring numbers, heritability and restrictions. For the scenarios examined, benefits of index selection exceeded 5% only for: low $h^2_{(0)}$ and large N, in which optimum indices placed substantial extra weight on family information compared with mass selection; and high $h^2_{(0)}$ and small N, in which optimum indices placed substantially less weight on the family than mass selection. As restrictions on ΔF become less severe or time horizons become shorter, it would be anticipated that the range of $h^2_{(0)}$ and N for which index selection is beneficial from putting extra weight on the family would expand, whereas the range benefiting from reduced weight on the family would diminish.

The study has re-evaluated and generalized the results of Robertson (1960) and Dempfle (1975) on the selection limits of indices. Robertson (1960) concluded that a selection proportion of 0·5 was optimum for mass selection and that mass selection was always superior to family selection (b_2, $b_3 \gg b_1$). Dempfle (1975) concluded that the selection limit may be greater for within-family selection than for mass selection, and that this was particularly evident when the accuracy of selection was high. In Dempfle's study the squared accuracy is akin to the heritability in this study, since he considered evaluation of the genotype as a unit (e.g. from progeny testing or phenotype) and not as a composite of separate bits of information on ancestors and a Mendelian sampling term.

Results of this study show optimum selection proportions for maximizing the selection limit to be more than 0·5, and equal in both sexes. The discrepancy arises because Robertson used Wright's formula for predicting ΔF, but this underestimates the impact of selection intensity upon rate of inbreeding (Woolliams *et al.*, 1993) and so appropriate modification tends to favour lower selection intensities. Although the infinitesimal model is unrealistic, particularly when considering responses at the limit, it is still a standard model and it is useful as a reference for comparison.

The optimum weights for maximizing the selection limit were close to mass selection with a greater emphasis on family information for low $h^2_{(0)}$ and less for high $h^2_{(0)}$. The results of Dempfle in comparing mass selection with within-family selection are therefore consistent with the results of this study; however, within family selection is not the optimum, and some positive weight should be given to family information (i.e. b_2, $b_3 > 0$). The authors find it remarkable that in maximizing the selection limit the optimum selection proportion was independent of heritability, and the optimum weights were close to mass selection.

When ΔF was restricted and in the absence of common environmental variance, the optimum index selection scheme had: more intense selection on sires and less emphasis on family information as $h^2_{(0)}$ increased up to 0·6; and more intense selection on sires and more emphasis on family information as the severity of the restriction on ΔF was reduced and more resources were available. With restrictions on ΔF, time horizons $t = 5$ or $t = 20$ made only minor modifications to optimum schemes with a small reduction in intensity (Table 2) and a slightly greater emphasis on family information for longer time horizons (Table 3).

In hierarchical schemes with $\sigma^2_C = 0$ the changes in selection intensity for females in response to variation in scheme parameters were not as smooth as for males, since the optimum mating ratio (which only takes discrete values) changed when selection on males was most intense; i.e. when $h^2_{(0)}$ was at the upper end of the range considered, resources were greater and restrictions on ΔF were less severe. This pattern was also noted for mass selection (Villanueva *et al.*, 1996) and differs from the earlier conclusion of Jódar & López-Fanjul (1977) who predicted maximum gain at any generation with $d = 1$. The discrepancy again lies in the adequacy of the prediction of ΔF. In the latter paper, Wright's formula was used, whereas better predictions, even one-generation predictions (Wray *et al.*, 1990), show that ΔF will depend in part on terms such as $i^2\phi d^{-1}N_s^{-1}$, where ϕ is the correlation between index values of either paternal half-sibs or full-sibs. Villanueva *et al.* (1996) argued that such terms under the conditions noted above are capable of favouring schemes with increased d. Schemes which are capable of fully factorial mating with equal information on both sexes will have a different outcome since there will also be terms in both $N_s N_d^{-1}$ and $N_d N_s^{-1}$. The complete symmetry of such schemes will lead to optima which are square designs with $N_s = N_d$ (Woolliams, 1989; De Boer & Van Arendonk, 1994).

The presence of common environmental variation increased selection intensity on males, increased the mating ratio for low heritabilities and generally decreased the emphasis on sib information, particularly full-sibs. These trends in the optimal schemes are predictable as they move to the extremes of the range of heritability presented. The value of information from half-sibs can be increased if the common environmental variation is averaged out over full-sib families, whereas the information from a full-sib family is always formally confounded with the common environment. Therefore there is a much greater pressure for d to increase above 1 when $\sigma^2_C > 0$ (with $b_3 > b_2$) than when $\sigma^2_C = 0$. A further influence

on the results for $\sigma_C^2 > 0$ that becomes increasingly important as $h_{(0)}^2$ increases and more weight is attached to within-family deviations is that a greater proportion of the variance of within-family deviations is associated with genetic variance than in the case with $\sigma_C^2 = 0$. Thus for the same heritability, the presence of substantial common environmental variance can increase gains in optimum indices under restricted inbreeding.

The formulae used are approximations, but where direct comparison with simulation has been made, excellent agreement was obtained. Further improvements are available. For example, the selection intensity used in the formulae was that appropriate for infinite populations with uncorrelated estimates of breeding values, which is a serious overestimate in small schemes when there are very high correlations in the indices among family members (Hill, 1976; Rawlings, 1976). However, in our case the restrictions on ΔF (even for $\lambda = 0$ when $t = 20$) increased the numbers selected and very much reduced the correlations among family members compared with classical weights (see Table 3) and consequently the impact on selection intensity is small. Applying the approximation of Meuwissen (1991) for $\lambda = 0$, $t = 20$, $h_{(0)}^2 = 0{\cdot}1$ and $N = 200$, showed the selection intensity was reduced from 1·346 to 1·324. The bias will be more important when maximizing $\Delta G_{(5)}$ for $\lambda = 0$, but corrections can easily be incorporated into the optimization procedure.

Index weights have been assumed constant over generations. However, the similarity in the optimum weights obtained when maximizing response at different generations under specific constraints on ΔF (Table 3) suggests that little improvement would be made by allowing the weights to change each generation.

The optimization procedure also makes the assumption where necessary of non-integer numbers of scored individuals per family and has not accounted for variation in family size in the formulae. However, the deviation of the schemes from a constant-integer family size for the parameters used has only a small impact. The worst case (that with the greatest coefficient of variation) was for $h_{(0)}^2 = 0{\cdot}1$, $N = 200$, $\Delta F \leqslant 0{\cdot}25\%$ where $N_s = N_d = 69$. Here, with 62 families with 3 offspring and 7 with 2 offspring the coefficient of variation of family size (CV) is 0·1. For $d = 1$, the rate of inbreeding is related to $E(n^2) = [E(n)]^2[1 + CV^2]$, where n is the family size, but here we have only considered $[E(n)]^2$. The proportional errors in ΔF introduced by neglecting variation in family size are therefore at most 1% (i.e. CV^2). Other factors such as reproductive limitations may also create variation in family size. Complications of index definition arise in practice if the coefficient of variation in the number of scored individual is large, since indices are no longer identical or uniformly accurate across families.

The study has optimized numbers of parents of both sexes and index weights pre-determined for constraints on total number of offspring per generation, rate of inbreeding and time horizon. The first two constraints represent resources available and the risk attached to the scheme. The latter would reflect the objectives of the scheme, where for example a competitive breeding company may have a short horizon and populations conserved *in situ* would have long horizons. Other restrictions, for example on the expected number scored per family arising from biological constraints on family size, could be added to the framework. The results have practical significance since they show that even when no restrictions are placed on the rate of inbreeding, optimal weights show substantial deviations from classical weights.

Appendix A. Calculation of ΔF_E

Wray *et al.* (1994) give expressions for the correlations of the index values of full-sibs (ρ_D) and half-sibs (ρ_H) for indices of the form used here. These are given by

$$\rho_D = \left\{ -b_1^2 \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} + \left[b_2^2 \left(1 - \frac{N_s}{N_d}\right) \times \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right] + b_3^2 \left[\sigma_{As}^2 + \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right] \right\} \sigma_I^{-2}.$$

$$\rho_H = \left\{ -b_2^2 \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] + b_3^2 \left[\sigma_{As}^2 + \frac{N_s}{N_d} \left[\sigma_{Ad}^2 + \sigma_C^2 + \frac{\sigma_{Aw}^2 + \sigma_E^2}{n} \right] \right] \right\} \sigma_I^{-2}.$$

The dependence on t has been left implicit above, but otherwise these have an identical form to equations 3 and 4 of Wray *et al.* (1994). In the calculation of ΔF_E the values after a single generation of selection were used for σ_{As}^2 and σ_{Ad}^2 while σ_{Aw}^2 was assumed to be constant.

Appendix 6 of Woolliams *et al.* (1993) uses a result of Mendell & Elston (1974) to show that

$$\text{Prob}(i,j \text{ full-sibs} \mid i \text{ of sex } x, j \text{ of sex } y, \text{ both selected}) = [(1/2)n - \delta_{xy}][(1/2)N - \delta_{xy}]^{-1} Q_D p_x^{-1},$$

where $Q_D = \Psi[(i_x \rho_D - \nu_y)(1 - k_x \rho_D)^{-1/2}]$ and Ψ denotes the cumulative normal distribution with zero mean and unit variance; i_x, p_x and k_x are the intensity of selection, proportion selected and variance reduction parameter for selection on sex x, and ν_y is the truncation deviate for selection on sex y; $\delta_{xy} = 1$ if $x = y$ (i.e. i and j same sex) and 0 otherwise and accounts for sampling without replacement. Although the probability is symmetric in x and y, the expression

Table B 1. *Predicted rates of inbreeding and percentage errors of predictions*

		$h^2_{(0)}$							
N_d	n	0·0		0·1		0·4		0·99	
20	6	2·47	(+1·6)	2·70	(+3·4)	2·34	(+2·2)	1·39	(+1·5)
40	6	2·30	(+0·9)	2·49	(+3·3)	2·03	(+2·0)	1·10	(+1·9)
	12	4·63	(+3·8)	4·77	(+6·2)	3·25	(+2·2)	1·26	(+2·4)
200	6	3·03	(+3·4)	2·78	(+8·2)	1·77	(+0·6)	0·83	(−1·2)
	12	4·64	(+2·7)	3·88	(+5·4)	2·34	(−0·8)	0·89	(+1·1)

is not: both forms are approximations but the use of $x = m$, $y = f$ is found to be more accurate assuming $p_m < p_f$. Similarly,

Prob (i,j half-sibs | i of sex x, j of sex y, both selected)

$$= [(1/2)n - \delta_{xy}][(1/2)N - \delta_{xy}]^{-1} Q_H p_x^{-1},$$

where $Q_H = \Psi[(i_x \rho_H - \nu_y)(1 - k_x \rho_H)^{-1/2}]$.

Problems of predicting co-selection probabilities were encountered by Wray *et al.* (1994) for when both (i) the number of males in a half-sib family was large compared with the number selected ($2N_s < N$ in this notation) and (ii) ρ_H was high. This circumstance was rarely encountered in this study because optimum restricted indices placed less emphasis on family information. Where necessary the solution adopted by Wray *et al.* (1990) was used, in which p_m was replaced by $p'_m = (1 - \rho_H) p_m + \rho_H N_s^{-1}$ for calculating the associate parameters i', k' and ν'.

Wray *et al.* (1990) derived a transition matrix of the form

$$\begin{pmatrix} r_1 & 1/2 & r_2 & (1/2)(1-2r_1-2r_2) \\ r_3 & 1/2 & r_4 & (1/2)(1-2r_3-2r_4) \\ r_4 & 1/2 & r_6 & (1/2)(1-2r_5-2r_6) \\ 0 & 1/2 & 0 & 0 \end{pmatrix}$$

where $r_1 = (1/4)\,\text{Prob}(i,j$ have distinct sires $| \, i \neq j$, both males, both selected); $r_2 = (1/4)\,\text{Prob}(i,j$ have distinct dams $| \, i \neq j$, both males, both selected). The terms r_3 and r_5 are defined as r_1 but considering different-sex and female-only pairs respectively; and r_4 and r_6 are defined as r_2 for different-sex and female-only pairs respectively. The terms r_1 and r_2 are

$r_1 = (1/4)$
$\times [1 - \text{Prob}(i,j$ half-sibs $| \, i, j$ both males, both selected)
$- \text{Prob}(i,j$ full-sibs $| \, i,j$ both males, both selected)]

$r_2 = (1/4)$
$\times [1 - \text{Prob}(i,j$ full-sibs $| \, i,j$ both males, both selected)]

and r_3, r_4, r_5 and r_6 can be constructed similarly using the probabilities derived above.

The estimate of ΔF_E which accounts for the co-selection in a single generation, but which does not account for the inheritance of selective advantage, is then calculated by obtaining the largest eigenvalue of the transition matrix. Wray *et al.* (1990) show the relationship of this estimate to a first-order approximation using variances of family sizes given by Hill (1979).

Appendix B. Validation of the method used to compute the rate of inbreeding under index selection

Table B 1 shows predicted rates of inbreeding and percentage errors of predictions (in brackets) calculated as $100 \times (P - S)/S$, where P and S represent predicted and simulated values, respectively, for schemes with $N_s = 20$ males and different numbers of females (N_d), numbers of offspring per mating (n) and heritabilities ($h^2_{(0)}$). Simulated values are from Wray *et al.* (1994).

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and by the European Union. S.A.C. 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 for useful comments on the manuscript.

References

Belonsky, G. M. & Kennedy, B. W. (1988). Selection on individual phenotype and best linear unbiased predictor of breeding value in a closed swine herd. *Journal of Animal Science* **66**, 1124–1131.

Brisbane, J. R. & Gibson, J. P. (1995). Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theoretical and Applied Genetics* **91**, 421–431.

Bulmer, M. G. (1971). The effect of selection on genetic variability. *American Naturalist* **105**, 201–211.

Caballero, A., Wei, M. & Hill, W. G. (1996). Survival rates of mutant genes under artificial selection using individual and family information. *Journal of Genetics* **75**, 63–80.

Cockerham, C. C. & Burrows, P. M. (1980). Selection limits and strategies. *Proceedings of the National Academy of Sciences of the USA* **77**, 546–549.

De Boer, I. J. M. & Van Arendonk, J. A. M. (1994). Additive response to selection adjusted for effects of inbreeding in a closed dairy cattle nucleus assuming a large number of gametes per female. *Animal Production* **58**, 173–180.

Dekkers, J. C. M. (1992). Asymptotic response to selection on best linear unbiased predictors of breeding values. *Animal Production* **54**, 351–360.

Dempfle, L. (1975). A note on increasing the limit of selection through selection within families. *Genetical Research (Cambridge)* **24**, 127–135.

Gomez-Raya, L. & Burnside, E. B. (1990*a*). The effect of repeated cycles of selection on genetic variance, heritability and response. *Theoretical and Applied Genetics* **79**, 568–574.

Gomez-Raya, L. & Burnside, E. B. (1990*b*). Optimum selection indices for continuous selection programmes. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XIII, Edinburgh, 23–27 July, pp. 273–276.

Grundy, B., Caballero, A., Santiago, E. & Hill, W. G. (1994). Use of biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. *Animal Production* **59**, 465–468.

Hill, W. G. (1976). Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32**, 889–902.

Hill, W. G. (1979). A note on effective population size with overlapping generations. *Genetics* **92**, 317–322.

Jódar, B. & López-Fanjul, C. (1977). Optimum proportions selected with unequal sex numbers. *Theoretical and Applied Genetics* **50**, 57–61.

Luo, Z. W., Woolliams, J. A. & Thompson, R. (1995). Controlling inbreeding in dairy MOET nucleus schemes. *Animal Science* **60**, 379–387.

Mendell, N. R. & Elston, R. C. (1974). Multifactorial qualitative traits: genetic analysis and prediction of recurrence risks. *Biometrics* **30**, 41–57.

Meuwissen, T. H. E. (1991). Reduction of selection differentials in finite populations with a nested full-half sib family structure. *Biometrics* **47**, 195–203.

Meuwissen, T. H. E. & Woolliams, J. A. (1994). Effective sizes of livestock populations to prevent a decline in fitness. *Theoretical and Applied Genetics* **89**, 1019–1026.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd edn, pp. 390–395. Cambridge: Cambridge University Press.

Quinton, M., Smith, C. & Goddard, M. E. (1992). Comparison of selection methods at the same level of inbreeding. *Journal of Animal Science* **70**, 1060–1067.

Rawlings, J. O. (1976). Order statistics for a special class of unequally correlated multinormal variates. *Biometrics* **32**, 875–887.

Robertson, A. (1960). A theory of limits in artificial selection. *Proceedings of the Royal Society of London, Series B* **153**, 234–249.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research (Cambridge)* **2**, 189–194.

Stam, P. (1980). The distribution of the fraction of the genome identical by descent in finite random mating populations. *Genetical Research (Cambridge)* **35**, 131–155.

The Numerical Algorithms Limited (1991). *NAG Fortran Library Manual*, Mark 15, 1st edn., Oxford: The Numerical Algorithms Ltd.

Toro, M. & Perez-Enciso, M. (1990). Optimization of selection response under restricted inbreeding. *Genetics, Selection, Evolution* **22**, 93–107.

Verrier, E., Colleau, J. J. & Foulley, J. L. (1990). Predicting cumulated response to directional selection in finite panmictic populations. *Theoretical and Applied Genetics* **78**, 833–840.

Verrier, E., Colleau, J. J. & Foulley, J. L. (1993). Long-term effects of selection based on animal model BLUP in a finite population. *Theoretical and Applied Genetics* **87**, 446–454.

Villanueva, B. & Kennedy, B. (1993). Index versus tandem selection after repeated generations of selection. *Theoretical and Applied Genetics* **85**, 706–712.

Villanueva, B., Woolliams, J. A. & Simm, G. (1994). Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics, Selection, Evolution* **26**, 517–535.

Villanueva, B., Woolliams, J. A. & Gjerde, B. (1996). Optimum designs for breeding programmes under mass selection with an application in fish breeding. *Animal Science* **63**, 563–576.

Wei, M., Caballero, A. & Hill, W. G. (1996). Selection response in finite populations. *Genetics* **144**, 1961–1974.

Woolliams, J. A. (1989). Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**, 1–14.

Woolliams, J. A. (1990). Strategies to maximise selection progress in dairy cattle. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XIV, Edinburgh, 23–27 July, pp. 15–24.

Woolliams, J. A. & Pong-Wong, R. (1995). Short- versus long-term responses in breeding schemes. In *46th Annual Meeting of the European Association for Animal Production*, Prague, 4–7 September.

Woolliams, J. A. & Thompson, R. (1994). A theory of genetic contributions. In *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, vol. 19, Guelph, 23–27 July, pp. 127–134.

Woolliams, J. A., Wray, N. R. & Thompson, R. (1993). Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research (Cambridge)* **62**, 231–242.

Wray, N. R. & Goddard, M. E. (1994). Increasing long term response to selection. *Genetics, Selection, Evolution* **26**, 431–451.

Wray, N. R. & Hill, W. G. (1989). Asymptotic rates of response from index selection. *Animal Production* **49**, 217–227.

Wray, N. R. & Simm, G. (1990). The use of embryo transfer to accelerate genetic improvement in beef cattle. In *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, vol. XV, Edinburgh, 23–27 July, pp. 315–318.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1990). Methods for predicting rates of inbreeding in selected populations. *Theoretical and Applied Genetics* **80**, 503–512.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1994). Prediction of rates of inbreeding in populations undergoing index selection. *Theoretical and Applied Genetics* **87**, 878–892.

Paper 27

EVALUATION OF EMBRYO SEXING AND CLONING IN DAIRY CATTLE NUCLEUS SCHEMES UNDER RESTRICTED INBREEDING

B. Villanueva[1], J.A. Woolliams[2] and G. Simm[1]

[1]Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, United Kingdom
[2] Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, United Kingdom

SUMMARY

The value of cloning and sexing of embryos for increasing rates of genetic progress under restricted inbreeding are examined in MOET dairy cattle schemes with progeny testing. Using a deterministic model, schemes are compared for maximum genetic gain at specific rates of inbreeding. Substantial increases in genetic gain are expected as a result of sexing and cloning of embryos when compared at equal rates of inbreeding. However, there is an intermediate optimum number of clones per genotype.

Keywords: Embryo cloning, embryo sexing, inbreeding

INTRODUCTION

Recent major advances in nuclear transfer from cultured cell lines (Campbell *et al.*, 1996) opens up the possibility of producing large numbers of genetically identical individuals (clones) in livestock species. Different studies investigating the value of cloning for accelerating genetic improvement of farm animals have given contradictory results depending on the specific restrictions and assumptions implied in the models (see review by Villanueva and Simm, 1994). This has been also the case when evaluating embryo sexing.

Although the primary concern over the use of clones is its impact on genetic variation, deterministic predictions of the benefit of this technique (Colleau, 1992) have focused solely on expected genetic gains. The objective of this study is to develop a deterministic model for maximising the rate of genetic progress while restricting the rate of inbreeding when using sexing and cloning of embryos. The value of these techniques is investigated in dairy cattle schemes with overlapping generations.

METHODS

The trait under selection is assumed to be controlled by an infinite number of loci with additive ($\sigma^2_A = 0.4$) and dominance ($\sigma^2_D = 0.2$) components and it is recorded only in females. Phenotypic variance (σ^2_P) is unity. Multiple Ovulation and Embryo Transfer (MOET) schemes with bull progeny testing are modelled. The total number of embryos transferred is t = 4000. Each year M = 100 are progeny tested outside the scheme on 50 daughters per bull and they are used for one year. Progeny proofs are available when males are six years old. Records on females correspond to their first lactation only and are available when females are three years old. Generations overlap with male and female pathways having lengths of seven

and four years respectively. The annual number of selected males (m) used for breeding (selected on their progeny test results) is optimised.

Two embryo collections ($r = 2$) are carried out at the beginning of the second lactation. Natural first and second calves are ignored (initial results showed that this simplification does not change the conclusions on cloning). A proportion $p = 0.7$ of the females respond to superovulation and the number of embryos per recovery in responders is $e = 5$. The embryo survival rate is $s = 0.6$ and the number of members per clone is $c = 1$, 3 or 5. If embryos are sexed only female embryos are cloned. The best f females are selected on their clonal means but only the best are used to breed the M males required for progeny testing. Therefore a number of dams (f_s) breed both sexes while the remaining selected females (f_r) are used to breed females only. Both f_s and f_r refer to distinct genotypes and are completely determined by reproductive constraints. They are given by $f_s = M / [(0.5)\ resc]$ and $f_r = [ts - (0.5) f_s\ resc\ (c + z)] / [(0.5)\ resc\ (c + y)]$ with $z = y = c$ when embryos are unsexed and $z = 1$ and $y = 0$ when embryos are sexed. Schemes without cloning are modelled by setting $c = 1$. Semen is assumed unsexed. Selected proportions for males (p_m), f_s females (p_s) and f_r females (p_r) are respectively $p_m = m/M$, $p_s = f_s / [(0.5) fresc]$ and $p_r = \{1/[(0.5)\ resc]\} - p_s$. Different males are used in different embryo recoveries.

The model determines the asymptotic rate of genetic progress (i.e. Bulmer equilibrium but ignoring changes in genetic variance due to inbreeding). Predicted asymptotic rates of genetic progress (ΔG) and inbreeding (ΔF) are obtained from long-term contributions theory (Woolliams, 1998). The annual rates of genetic gain and inbreeding are

$$\Delta G = m\ E(\mu_{i(m)}\alpha_{i(m)}) + f_s\ E(\mu_{i(s)}\ \alpha_{i(s)}) + f_r\ E(\mu_{i(r)}\alpha_{i(r)})$$

$$\Delta F = (1/2)[m\ E(\mu^2_{i(m)}) + f_s\ E(\mu^2_{i(s)}) + f_r\ E(\mu^2_{i(r)})]$$

where $\mu_{i(x)}$ is the expected long-term contribution from an ancestor i of breeding category x in generation t_1 to descendants in generation t_2 ($t_2 >> t_1$) and $\alpha_{i(x)}$ is its Mendelian sampling term. There are three categories of breeding animals (males, $x = m$, f_s females, $x = s$, and f_r females, $x = r$). The expected long-term contribution conditional on breeding value for each category can be expressed as $\mu_{i(x)} = (a_{(x)} + b_{(x)} A_{i(x)}) / n_{(x)}$, where $n_{(x)} = m, f_s$ or f_r according to category and $A_{i(x)}$ is the breeding value of the ancestor (deviated from the mean of the selected individuals in that category). The vector a is the right eigenvector of $\mathbf{G}^T$ and

$$\mathbf{b} = [\mathbf{I} - (\mathbf{G}^T \otimes \mathbf{B}^T)]^{-1} (\mathbf{G}^T \otimes \mathbf{N}^T)\mathbf{a}$$

Here **G** is the matrix describing the expected dispersion of genes through the population accounting for selection between categories, **B** is the matrix of regression coefficients of the breeding values of selected offspring on the breeding value of the parent and **N** is the matrix of regression coefficients of the number of selected offspring on the breeding value of the parent. For ΔF, Poisson family sizes are assumed and $\mu_{i(x)}$ are completed by addition of terms

of the form $b_{(y)}A_{i(y)}$ to account for breeding values of randomly assigned mates.

With mass selection the number of sires to be selected (m) is the only variable to be optimised. The optimum value for m is that which maximises ΔG with an upper bound on ΔF. The value for m is increased until the constraint on ΔF is satisfied.

RESULTS

Without embryo cloning (c = 1) the rate of genetic gain was higher (up to 3%) in schemes with sexed embryos than in schemes without sexing (Table 1). Without sexing both male and female embryos need to be transferred to recipients (although some male offspring are later discarded). The unnecessary transfer of males in schemes with unsexed embryos led to lower selection differential than in schemes with sexing (Table 2).

Table 1. Annual rate of genetic progress (ΔG; σ_P units) achieved with optimum number of sires (m) in schemes with and without embryo sexing and cloning when the rate of inbreeding per generation (ΔF) is constrained to two specific values. The numbers of f_s and f_r dams are determined by reproductive parameters

				$\Delta F \leq 1\%$		$\Delta F \leq 0.5\%$	
Sexing	cloning	f_s	f_r	m	ΔG	m	ΔG
No	$c = 1$	33	367	4	0.135	7	0.126
	$c = 3$	11	33	5	0.147	12	0.131
	$c = 5$	7	9	8	0.144	44	0.108
Yes	$c = 1$	33	733	4	0.139	7	0.129
	$c = 3$	11	74	5	0.150	11	0.136
	$c = 5$	7	24	6	0.153	23	0.127

Table 2. Selection intensities (i) and accuracies (ρ) for optimum schemes to maximise the annual rate of genetic progress when the rate of inbreeding per generation is constrained to 1% with and without embryo sexing and cloning

Sexing	cloning	i_m	i_s	i_r	ρ_m	ρ_f
No	$c = 1$	2.153	2.296	0.982	0.889	0.562
	$c = 3$	2.062	2.296	1.507	0.886	0.668
	$c = 5$	1.858	2.296	1.684	0.885	0.698
Yes	$c = 1$	2.153	2.534	1.025	0.889	0.562
	$c = 3$	2.062	2.534	1.580	0.886	0.667
	$c = 5$	1.984	2.534	1.773	0.885	0.697

Substantial extra gains were also obtained with embryo cloning for moderate values of c due to increased female selection accuracy (Tables 1 and 2). The highest increase in ΔG found from embryo cloning was around 10% (sexing, $c = 5$, $\Delta F \leq 1\%$). There was an optimum value for c for maximising response under restricted inbreeding. The number of dams is greatly decreased when embryos are cloned. As a consequence of this the number of sires needs to be increased to meet the constraint on ΔF. Thus, when ΔF is restricted and c is high the decrease in male selection intensity does not compensate for the increase in female accuracy. The optimum c decreased with the severity of the restriction applied to ΔF.

DISCUSSION

Colleau (1991) found a benefit from embryo sexing in MOET schemes with progeny testing in dairy cattle with restrictions on the number of bulls progeny tested, the total number of embryos transferred and the number of sires. Here we have extended Colleau's model to optimise the number of sires for maximising gain when the inbreeding rate is restricted to specific values. As in Colleau's (1991) study, we found a similar magnitude of benefit from sexing when $\Delta F \leq 1\%$ because the number of sires he assumed was close to the optimum for the restriction applied to ΔF.

Cloning of embryos has been shown also to increase the rate of genetic gain in MOET dairy cattle schemes with progeny testing and restricted inbreeding. The extra gain was higher than that found in schemes without progeny testing (de Boer *et al.*, 1994) indicating the advantage of the extra male accuracy obtained outside the scheme. With unrestricted inbreeding and fixed number of males selected the rate of progress increases with the number of members per clone (Colleau, 1992). However, when the number of sires is optimised to maximise gain under a specific constraint on ΔF then there is an optimum value for the number of replicates per clone. The procedure used here for finding the optimum use of cloning in the breeding population can be extended to selection on an index using family information. Nonetheless, the optimum use of cloning for disseminating superior genotypes to the commercial population still remains unknown.

ACKNOWLEDGEMENTS

This work was funded by the BBSRC, EU, SOAEFD, MMB, MLC and MAFF.

REFERENCES

Campbell, K.H.S., McWhir, J., Ritchie, W.A. and Wilmut, I. (1996) *Nature* **380**:64-66.

Colleau, J.J. (1991) *J. Dairy Sci.* 74:3973-3984.

Colleau, J.J. (1992) *Genet. Sel. Evol.* 24:345-361.

De Boer, I.J.M., Meuwissen T.H.E. and van Arendonk A.M. (1994) *Anim. Prod.* 59: 345-358.

Villanueva, B. and Simm, G. (1994). *Proc. 5th World Congr. Genet. Appl. Liv. Prod., Vol 20.* Guelph, 23-27 July, pp 200-207.

Woolliams, J.A. (1998). *Proc. 6th World Congr. Genet. Appl. Liv. Prod.*. Armidale, 11-16 January.

Paper 28

T. H. E. Meuwissen · J. A. Woolliams

Effective sizes of livestock populations to prevent a decline in fitness

Received: 27 January 1994 / Accepted: 6 July 1994

Abstract In livestock populations, fitness may decrease due to inbreeding depression or as a negatively correlated response to artificial selection. On the other hand, fitness may increase due to natural selection. In the absence of a correlated response due to artificial selection, the critical population size at which the increase due to natural selection and the decrease due to inbreeding depression balance each other is approximately $D/2\sigma^2_{wa}$, where D=the inbreeding depression of fitness with complete inbreeding, and σ^2_{wa}=the additive genetic variance of fitness. This simple expression agrees well with results from transmission probability matrix methods. If fitness declines as a correlated negative response to artificial selection, then a large increase in the critical effective population size is needed. However, if the negative response is larger than the response to natural selection, a reduction in fitness cannot be prevented. From these results it is concluded that a negative correlation between artificial and natural selection should be avoided. Effective sizes to prevent a decline in fitness are usually larger than those which maximize genetic gain of overall efficiency, i.e., the former is a more stringent restriction on effective size. In the examples presented, effective sizes ranged from 31 to 250 animals per generation.

Key words Critical effective population size
Inbreeding depression · Natural selection · Fitness
Conservational biology

Communicated by J. S. F. Barker

T. H. E. Meuwissen (✉)
Institute for Animal Science and Health,
P.O. Box 501,3700 AM Zeist,
The Netherlands

J. A. Woolliams
Institute for Animal Physiology and Genetics Research,
Roslin, Midlothian EH25 9PS,
Scotland, UK

Introduction

New developments in animal breeding schemes are designed to increase genetic gain, but inbreeding rates are often increased concomitantly. For example, the introduction of BLUP as a means of estimating breeding values (Henderson 1984) and of MOET nucleus schemes (Nicholas and Smith 1983) will increase annual rates of inbreeding because (1) the weightings attached to pedigree information are increased, (2) generation intervals are decreased, and (3) because of selection within small nucleus populations in the case of MOET schemes. The question arises, what rate of inbreeding is justifiable in a breeding scheme. Restricting rates of inbreeding may have a major impact on the optimization of breeding schemes. For instance, mass selection can be superior to BLUP selection when the constraint on inbreeding is severe (Quinton et al. 1992).

Detrimental effects of inbreeding are: (1) the reduction of additive genetic variance, which reduces rates of response and limits to selection for the trait under selection and other traits; (2) inbreeding depression for the trait under selection, if gene effects are non-additive; (3) inbreeding depression in fitness of the animal. However, in his simulation of dairy cattle breeding, Meuwissen (1989) found that reduction in additive genetic variance, over ten generations of selection, would only reduce total genetic gain by 13%. It was assumed, that the effective population size (N_e) was small (ten animals), the accuracy of selection was high (about 0.8), and the proportion selected was small (0.1). These assumptions reduce cumulative genetic gain more than is likely in practice. In this rather extreme situation, doubling the N_e increases the rate of gain by approximately 6.5%.

Because most production traits, such as growth rate and feed efficiency, do not show a high degree of inbreeding depression (Falconer 1981), inbreeding depression is not a major factor for production traits. In contrast, fitness, a combination of survival and reproductive traits, is often not selected for directly but is highly influenced by inbreeding depression (e.g., Beilharz 1982; Wiener et al.

1992c). Hence, in the medium term, the most stringent restriction on the rate of inbreeding is imposed by inbreeding depression of fitness.

Rates of deterioration of fitness traits through selection were derived by Robertson (1966). These rates may be ameliorated by incorporating them in selection indices, e.g., by desired gains indices (Cunningham et al. 1970). This strategy requires the measurement of components of fitness and knowledge of their relative importance for fitness and of their genetic parameters. On the other hand, N_e may be chosen such that depression of fitness due to inbreeding equals the additive genetic improvement due to natural selection, so that fitness should not change. Thus, Soulé (1980) considers natural selection as the remedy against the fixation of deleterious alleles, while Franklin (1980) and Soulé (1980) provide crude empirical values of minimal viable effective population sizes.

The aim of the present paper is to assess the critical effective population size for which decrease of fitness due to inbreeding depression and increase of fitness due to natural selection balance one another. The perspective of animal breeders is rather short term so that the accumulation of new mutations is ignored. Accumulation of mutations affects the long-term genetic variance and thus the response (Hill 1982) and the time till extinction of populations due to mutational meltdown (Lynch and Gabriel 1990). A simple formula, which ignores the effects of selection on genetic variances (Bulmer 1971) is derived. More complicated models are compared to the simple formula.

Materials and methods

A population with discrete generations is assumed which is subject to artificial selection. Not all animals have maximum fitness, so that variation for fitness exists, which is partly due to additive and partly due to dominant gene effects. Epistatic gene interactions are assumed to be absent. Let N_m be the total number of male parents with n_{mi}, the number of offspring produced by the ith male parent. Fitness is defined as $w_{mi}=n_{mi}/\bar{n}_m$, where $\bar{n}_m=\Sigma n_{mi}/N_m$ is the average number of offspring per male parent. Note that average fitness is $\Sigma w_{mi}/N_m=1$. Fitness of female parents is defined similarly. Thus, artificial selection precedes natural selection within a generation and fitness is treated as a maternal/paternal trait. If production traits undergoing artificial selection are genetically uncorrelated to fitness then gene frequencies and genetic variances are expected to remain unchanged by artificial selection (except as influenced by effective population size which is affected by intensity of selection).

A simple formula for critical effective population size

The fundamental theorem of Fisher (1929) states that the expected increase in fitness through natural selection on viability is σ^2_{wa} per generation, where σ^2_{wa} is the additive genetic variance of fitness. This conclusion may be drawn from the following argument. The phenotypic selection differential of male parents is the fitness of male parents minus the population mean of 1 weighted by their number of offspring:

$$X_{Sm}=\Sigma(w_{mi}-1)w_{mi}/N_m=\Sigma(w_{mi}-1)^2/N_m+\Sigma(w_{mi}-1)/N_m=V(w_m).$$

The same holds for female parents, so that $X_{Sf}=V(w_f)$.

If the mating of parents is at random with respect to fitness, then, from regression theory, the increase in breeding value for fitness is given by $\frac{1}{2}[X_{Sm}\ V(w_{am})/V(w_m)+X_{Sf}\ V(w_{af})/V(w_f)]=\sigma^2_{wa}$, where $V(w_{ax})$=the additive genetic variance of fitness in sex x, and $\sigma^2_{wa}=\frac{1}{2}[V(w_{am})+V(w_{af})]$. Thus, the expected increase in relative fitness through natural selection is σ^2_{wa} per generation. This derivation invokes *linear* regression of genotypes on phenotypes, which practically implies multi-normality, i.e., the infinitesimal model and normal distributed environmental effects. However, because selection differentials due to natural selection are small, predictions from linear regression will hold approximately even if distributions are non-normal and do not exhibit linear regression, e.g., in the case of genetic models with few loci. Jacquard (1972) shows that the theorem is approximately correct even for the one-locus model.

In generation t, the increase of inbreeding depression is $(F_t-F_{t-1})D$, where F_t is the inbreeding coefficient at generation t and D is the depression of fitness in percentages of the mean per percent of inbreeding. If ΔF is the rate of inbreeding as defined by Falconer (1981), this increase per generation is $\Delta F(1-F_{t-1})D$. With an additive genetic model, increase in fitness due to natural selection in generation t is $(1-F_{t-1})\sigma^2_{wa}$. Hence, two forces, natural selection increasing fitness and inbreeding depression reducing fitness, balance if $\sigma^2_{wa}=\Delta F\ D$, i.e., if effective population size is $N_e=D/2\sigma^2_{wa}$, because $\Delta F=1/2N_e$. But, if inbreeding depression occurs, the genetic model clearly includes dominance.

The following argument suggests that, even with dominance, an effective population size of $D/2\sigma^2_{wa}$ would prevent a decline of fitness in later generations. Increase of fitness due to natural selection, as derived above, is the sum of the covariances between sire and offspring and that between dam and offspring. The covariance between parent and offspring is derived for a model that includes dominance and inbreeding in Appendix 1. Appendix 1 shows, that for traits with low heritabilities, high coefficients of variation, and high inbreeding depression, which is generally the case for fitness traits, the covariance between parent and offspring exceeds $\frac{1}{2}(1-F_t)\sigma^2_{wa}$, where again ΔF is assumed to be small. Hence, increase in fitness would exceed $\sigma^2_{wa}(1-F_t)$ and $D/2\sigma^2_{wa}$ is a conservative estimate of the critical effective population size.

The two forces, inbreeding depression decreasing average fitness and natural selection improving additive genetic value, are assumed to be additive. In view of the complex nature of the stochastic process of changing gene frequencies, this approximation will be tested against results from transmission probability matrices describing the evolution of gene frequency distributions (Narain and Robertson 1969).

Let A_1 and A_2 denote the positive and negative allele for fitness, respectively. In generation t, the frequency of A_2 is q_t. Further the relative fitnesses of genotypes A_1A_1, A_1A_2 and A_2A_2 are (1+a), (1+d), and (1−a), respectively. Gametes of selected animals are assumed to unite at random, hence, the frequencies of the genotypes are $(1-q_t)^2$, $2q_t(1-q_t)$, and q_t^2, respectively. The frequency of A_2 after selection is $s^*(q_t)=[q_t^2(1-a)+q_t(1-q_t)(1+d)]/[q_t^2(1-a)+2q_t(1-q_t)\cdot(1+d)+(1-q_t)^2(1+a)]$. The denominator is the mean fitness.

A monoecious population of size N is assumed which has a probability of selfing of 1/N, and hence $N=N_e$. Let p_{it} be the probability that the frequency of A_2 is i/2N at generation t, for i=0...,2N. The vector p_t contains the elements p_{it}. If $q_t=j/2N$ in generation t, $q_{t+1}=i/2N$ with probability T_{ij}. From the binomial distribution:

$$T_{ij}=\binom{2N}{i}[s^*(j/2N)]^i\,[1-s^*(j/2N)]^{2N-i}.$$

If T denotes the matrix with elements T_{ij}, the evolution of the distribution of gene frequencies is given by the recurrence relationship: $p_{t+1}=Tp_t$ (Narain and Robertson 1969) from which the evolution of the distribution of gene frequencies can be obtained given an initial distribution; here: $q_0=0.2$ will be used.

Correlation between fitness and production and variance reductions

If a correlation exists between fitness and production, then here this correlation will be assumed to be negative. This assumption is made

because (1) only negative correlations cause a problem if decline in fitness is to be prevented, and (2) negative correlations are more likely as the resources of an animal are limited and merit in one trait may be (partly) offset by demerit in another: hence, the assumption that selection for production will most likely lead to a correlated decrease in fitness. Because selection for production precedes selection for fitness, reduction in σ^2_{wa} due to selection will need to be accounted for. The $1-F_t$ terms in the expressions for reduction in variance and inbreeding depression cancel as shown in the previous section. The non-additive genetic terms will be ignored which will lead to conservative estimates of response to natural selection (see Appendix 1).

In the following, genotypes for fitness and production will be assumed to involve many alleles such that the genotypic distributions are approximately normal. Let $C(i,j)_t$ be the (co)variance between trait i and j in generation t, where i (j) is P_p, P_a, W_p, or W_a, which are the phenotypes and additive genotypes of production and of fitness, respectively (if i=j, the variance of i is denoted). Further, * denotes after selection for production. For simplicity, sires and dams are assumed to be identically selected. Extension to differential sire and dam selection is straightforward. From regression theory (Pearson 1903):

$$C(i,j)_t^* = C(i,j)_t - C(i,P_p)_t\, C(j,P_p)_t\, k_p / C(P_p,P_p)_t,$$

where k_p is the reduction in variance of $C(P_p,P_p)_t$. For truncation selection $k_p = i_p(i_p - x_p)$, where i_p and x_p are the standardised selection differential and truncation point, respectively. The correlated response of fitness is:

$$\Delta w_{pt} = i_p\, C(P_p,W_a)_t / C(P_p,P_p)_t.$$

Let $\bar{w}$ denote the average fitness after selection for production:

$$\bar{w} = 1 + \Delta w_{pt}.$$

The selection differential of fitness due to natural selection is obtained by weighting the fitness of the selected animals minus the mean of the selected animals by their number of offspring:

$$\Sigma w_i(w_i - \bar{w})/(\bar{w}\, N) = \Sigma(w_i - \bar{w})^2/(\bar{w}\, N) + \Sigma(w_i - \bar{w})/N = C(W_p,W_p)_t^*/\bar{w}.$$

The regression coefficient of W_a on W_p is $C(W_a,W_p)_t^*/C(W_p,W_p)_t^*$. The genetic response due to natural selection equals the selection differential times the regression coefficient:

$$\Delta w_{wt} = C(W_a,W_p)_t^*/\bar{w}.$$

Let $C(i,j)_t^{**}$ denote (co)variances of parents weighted by their number of offspring, because these (co)variances determine the (co)variances due to parents observed among the offspring. Again applying regression theory:

$$C(i,j)_t^{**} = C(i,j)_t^* - C(i,W_p)_t^*\, C(j,W_p)_t^*\, k_{wt} / C(W_p,W_p)_t^*.$$

As shown in Appendix 2, $k_{wt} = C(W_p,W_p)_t^*/\bar{w}^2$.

With identical male and female selection, the additive genetic (co)variances in the next generation are (Bulmer 1971),

$$C(i,j)_{t+1} = \tfrac{1}{2}C(i,j)_t^{**} + \tfrac{1}{2}C(i,j)_0,$$

where i (j) denote the additive genotypes of production or fitness. The second term represents variance due to Mendelian sampling. The phenotypic variances and covariances are obtained by:

$$C(r,s)_{t+1} = C(i,j)_{t+1} + C_E(r,s),$$

where r(s) denotes the phenotype corresponding to genotype i(j), and $C_E(r,s)$ is the environmental (co)variance of traits r and s.

This model will be used to investigate effects of variance reduction due to selection and correlated responses from artificial selection on the effective population size required to prevent fitness deterioration. Inbreeding depression and correlated response from artificial selection will decrease genetic values for fitness. Natural selection will increase fitness. The critical effective population at which these effects balance will be compared to that predicted by $D/2\sigma^2_{wa}$.

Results

Comparing $D/2\sigma^2_{wa}$ to results from transmission probability matrix methods

Table 1 compares the population numbers for which the mean fitness is maintained at the initial level after ten generations of natural selection, using the transmission probability method. These results are compared to the prediction $D/2\sigma^2_{wa} = D/2h^2CV^2$. The ranges of coefficients of variation and heritabilities of fitness in Table 1 agree with the coefficients of variation and heritabilities of egg production in poultry and litter size in pigs and sheep (Smith 1984). With a coefficient of variation of 0.4, a fitness of w=0 is 2.5 standard deviations below the mean. Hence, a small proportion of the population would fail to produce offspring.

The values of gene effects a and d follow from the assumptions D=1, σ^2_{wa} (which is given in Table 1 as h^2CV^2), the initial gene frequency q_0=0.2, and the number of loci, L. The value of d is obtained as

$$d = D/2Lq_0(1-q_0), \qquad (1)$$

and that of a as

$$\sigma_{wa}^2 = 2Lq_0(1-q_0)[a + d(2q_0-1)]^2, \qquad (2)$$

which yields a quadratic in a with only one positive solution.

Table 1 The minimum effective population sizes to maintain fitness at its current level as derived from transmission probability matrices (N_e) compared with $D/2h^2CV^2$, where D=1 is the depression with complete inbreeding for different values of the positive homozygote (a) and the heterozygote genotype (d) at one locus. All loci are assumed to have equal a and d-values and initial gene frequencies of q_0=0.2

CV	h^2	No. of loci								
		20		40		80		160		
		a (×100)	N_e	a (×100)	N_e	a (×100)	N_e	a (×100)	N_e	$D/2h^2CV^2$
0.2	0.05	11.1	255	5.9	245	3.2	245	1.8	255	250
	0.10	11.9	120	6.5	115	3.6	120	2.1	125	125
0.4	0.05	12.9	55	7.2	55	4.1	60	2.4	60	63
	0.10	14.4	25	8.2	25	4.8	30	2.9	30	31
d (×100)		15.6		7.8		3.9		2.0		

Table 2 Critical effective population sizes (N_e) to maintain average fitness, when depression due to complete inbreeding D=1, phenotypic and genetic correlations between fitness and production are r, heritability and coefficient of variation of fitness are h_w^2 and CV, respectively, and the standardised selection differential and heritability of production, which is improved by mass selection, are i_p and 0.25, respectively. The results are proportional to D and are compared to $D/2\sigma_{wa}^2$

r	CV	h_w^2	N_e		$D/2\sigma_{wa}^2$
			i_p=1	i_p=2	
0.0	0.2	0.05	250	250	250
		0.10	126	126	125
	0.4	0.05	63	63	63
		0.10	32	32	31
−0.1	0.2	0.05	X[a]	X	250
		0.10	440	X	125
	0.4	0.05	130	>10000	63
		0.10	50	108	31
−0.2	0.2	0.05	X	X	250
		0.10	X	X	125
	0.4	0.05	X	X	63
		0.10	127	X	31

[a] X=the negative correlated response of artificial selection exceeds the natural selection response so that a decline in fitness cannot be prevented

The approximation $D/2\sigma_{wa}^2$ generally overestimates slightly the critical N_e (Table 1), as expected from Appendix 1. Overestimation is largest with small numbers of loci and with large σ_{wa}^2. Appendix 1 shows that in such situations contributions of non-additive terms to covariance between parent and offspring is largest. Increased covariance between parent and offspring implies more response from natural selection and, hence, a smaller critical population size.

If inbreeding depression is large and additive genetic variance is small, the effects of genes need to be overdominant to satisfy these conditions, i.e., d>a (see Table 1). This situation occurs particularly when the number of loci determining fitness is moderate or small. With overdominant gene action, selection tends to intermediate gene frequencies which are difficult to maintain, because drift leads to extreme gene frequencies, so that a large critical N_e is needed if gene effects are markedly overdominant.

Correlated response and reduction in variance due to selection

Table 2 provides the critical N_e after accounting for variance reduction due to selection and correlations between fitness and production. The results are for D=1, but results for different values of D can be obtained by multiplying the N_e of Table 2 by D.

If the correlations between fitness and production are zero, $N_e=D/2\sigma_{wa}^2$ is a good approximation for the critical N_e (Table 2). This indicates that reduction in variance due to natural selection is negligible, as expected for traits with low heritability. Natural selection cannot compensate for a negative correlated response of artificial selection, unless artificial selection is weak and the negative correlation is close to zero (Table 2). The critical N_e is markedly increased by a negative correlated response. If the correlated response exceeds the natural selection response, a decline of fitness (irrespective of the effective population size) cannot be prevented.

Discussion

The genetic model

Critical effective population sizes were obtained by balancing the effects of natural selection and inbreeding depression for a finite number of loci using a genetic model that encompassed non-additive genetic variation in the form of dominance but not epistasis. Depression was assumed to be proportional to the coefficient of inbreeding, which is a property of dominance but not generally of epistasis. Where this has been tested, dominance was the major cause of heterosis in between-line crosses but contributions of epistasis were significant for some traits (e.g., Abplanalph et al. 1984; Fairfull et al. 1987; Wiener et al. 1992a, b, c). The dominant effect of genes d is proportional to 1/L and the additive effect a is approximately proportional to $1/\sqrt{L}$ [see formulas (1) and (2)], hence, $\lim_{L\to\infty} d/a=0$, i.e., gene effects are additive. Consequently, the number of loci was assumed to be finite.

When considering the genetic background of fitness, the effects of very rare recessive deleterious (or even lethal) genes are most striking. But such genes do not cause much additive variation of fitness. Therefore, and because the problems caused by these genes may be more efficiently tackled by genetic markers, the genetic model used here was not directed in particular towards such genes. Genes with smaller detrimental effects on fitness are more likely to increase in frequency and collectively lead to a substantial reduction in fitness. Also non-linearity of the heritability of fitness (Frankham et al. 1988; Frankham 1990) was not considered here. An average heritability weighted by the genetic contributions of the selection candidates was used.

Prevention of deterioration of fitness

In the introduction, the effects of inbreeding on fitness were argued to be a more stringent restriction on population sizes of livestock than inbreeding effects on production traits. Goddard and Smith (1990) maximized the genetic gain of economic efficiency by optimizing the number of bull sires selected. Reduction in genetic variance and depression of efficiency due to inbreeding were both considered. Goddard and Smith concluded that the optimum number of bull

sires selected was ten bulls per generation, to be used equally (Poisson distribution of the number of offspring per bull). If Wright's (1931) formula, which ignores selection, is used and if the number of dams is infinite, an effective population size of 40 is obtained. This number is an overestimate, because the effect of selection on effective population size is large (Wray and Thompson 1990) and because the number of dams is finite. Further, it exceeds only the smallest effective sizes found in Table 2, i.e., those for CV=0.4 and h_w^2=0.1. Thus, preventing a decline in fitness would require more stringent restriction of effective population size than optimization of overall efficiency.

Goddard (1992) found an optimum number of bulls selected for the world-wide black and white cattle population of six per year, which is 30 per generation and an effective population size of 120 animals per generation using Wrights' (1931) formula. Because Wrights' formula overestimates effective population sizes substantially, this N_e remains smaller than the sizes presented in Table 1, except those in situations with high coefficients of variation.

Although some may doubt whether a change in fitness should be restricted to zero if selection is for overall efficiency there are reasons which make maintenance of fitness desirable. Clearly, the ability to survive and reproduce are vital, but some reduction may be compensated for by increased production. As fitness decreases its economic value will increase, which, ideally, will lead to a zero change at some critical value of fitness. Nevertheless, alowing fitness to decline involves several risks: valuable genes may be lost; critical values of components of fitness may be overshot due to the time lags in the selection process; and unfit animals may lead to public concern about the production system and, eventually, to the decreased use of products. The trend towards larger farm sizes calls for more 'trouble free' animals. Prevention of a decline now, rather than be forced to address reduced fitness at some point in the future, seems reasonable.

The assumption has been made that a population has a desirable level of fitness at present and that any deterioration would be undesirable. This assumption may not hold immediately for populations that have been developed with large effective population sizes. In the short term, such populations may be maintained at lower N_e than $N_e=D/2\sigma_{wa}^2$. If fitness is considered too low, then larger effective population sizes would give an increase in fitness of approximately $(1-F_t)(\sigma_{wa}^2-D/2N_e)$.

The values of D and σ_{wa}^2

Assuming D=1, as in Tables 1 and 2, means that fitness will decrease to zero as inbreeding approaches 1. Values of D for major components of fitness are available from the literature and are usually between 0.5 and 1% of inbreeding depression per percent of inbreeding (e.g., Falconer 1981, p. 228; Woodard et al. 1982; MacNeil et al. 1989; Wiener et al. 1992c). Because of non-linearity and because inbreeding rates are low, inbreeding depression will be expressed per percent of inbreeding, which is denoted by δ. Because overall fitness is the product of its components, for instance overall fitness may be survival rate times litter size, the depression of overall fitness is larger than that of its components:

$\delta=1-\Pi_{all\ j}(1-\delta_j)$, which is for a small δ_j equal to $\Sigma_{all\ j}\ \delta_j$,

where δ and δ_j are the depression of overall fitness and the jth component of fitness per percent of inbreeding, respectively. If some components of fitness are neglected in this formula, D will be underestimated. Latter and Robertson (1962) and Beilharz (1982) also show that depression of overall fitness is substantially larger than that of the individual components.

The additive genetic effect of overall fitness is the product of the additive genetic effects of its component traits. Using a Taylor series approximation, the additive genetic variance is approximated by:

$$\sigma_{wa}^2=\Sigma j\ CV_j^2h_j^2+\Sigma_j\Sigma_{ij}\ r_{ija}h_ih_jCV_jCVi,$$

where CV_j (h_j^2) is the coefficient of variation (heritability) of the jth component of fitness and r_{ija} is the additive genetic correlation between the ith and jth component. Neglecting components of fitness results in underestimation of both σ_{wa}^2 and D. Hence, the effect of neglecting components of fitness on the estimate of the critical N_e will depend on the specific situation.

The relationship between artificial and natural selection

It was assumed that high production could not compensate for low fitness and vice versa. In practical breeding schemes, natural selection response may be reduced by (1) greater management efforts to overcome low fitness of extreme producing animals, and (2) some standardization of family size if test places are limited. Whilst both these actions may tend to increase the critical effective population size required to maintain fitness, the effect of standardization of family size is more ambiguous because standardization also reduces the variance of family size and thus increases the effective population size.

Table 2 shows that natural selection can hardly compensate for a correlated response from artificial selection which decreases fitness. Hence, a negative correlation between fitness and selection indices should be prevented to avoid decreasing fitness, which can be achieved by using desired gains indices (Cunningham et al. 1970). Probably only the major components of fitness need to be included in the desired gains index. If some components, which have a negative correlation with the selection index are overlooked, they would become evident as selection proceeds and could then be included in the index. Natural selection will prevent deterioration of fitness components, which are uncorrelated to the selection index, if N_e is sufficiently large.

An implication of these considerations is the use of restricted indices. However Gibson and Kennedy (1990) suggest that restricted indices provide less genetic gain than

unrestricted indices. Their conclusion assumes that a decrease in fitness may be compensated by an increase of production and that all influences on profitability are clearly identified and accurately costed. Moreover, a large decrease in fitness will lead to losses due to infertile or inviable animals which cannot be compensated by production, i.e., the profit function is non-linear. Goddard (1983) showed that non-linear profit functions are optimized by searching for the maximal profit that can be reached within the time horizon of the breeding plan, and then selecting with linear indices towards this optimum. He also showed that if a long-term perspective is taken and the population is at the optimum for, e.g., fitness, which may be approximately the case at the onset of selection, the optimal selection strategy is to maintain fitness at its optimum.

In conclusion, a simple formula was derived for the critical effective population size at which natural selection for fitness and inbreeding depression balance. The results agree with results from transmission probability matrices for genetic models with dominance and overdominance (Table 1). The effective size to accomplish this goal is generally larger than that obtained from maximizing the genetic gain of economic efficiency. Further, negative correlated response for fitness from artificial selection increased the required critical effective sizes substantially, and in some circumstances this deleterious response could not be offset by increasing population size. Hence, an effective population size of $D/2\sigma^2_{wa}$ is recommended and the correlation between fitness and the selection indices for production should be prevented from being negative.

The critical effective population sizes presented apply also to natural populations, where the critical effective population size is the minimum population size such that natural selection can make progress against inbreeding depression. Populations which are smaller than the critical size will presumably go to extinction as they enter a downward spiral of ever-decreasing fitness. The results also apply to conservation biology by providing the minimum sizes required for conservation of populations.

Acknowledgements We are indebted to the reviewers for their useful contributions. The Milk Marketing Board of England and Wales, the Meat and Livestock Commission and the Ministry of Agriculture, Fisheries and Food are gratefully acknowledged for their sponsorship.

Appendix 1

The genetic covariance between parent and offspring with dominance and inbreeding

Loci will be assumed independent; hence, variances and covariances at individual loci sum to the variances and covariances of genotypes. This implies that linkage disequilibria due to selection and small population size are neglected. Weir and Cockerham (1974) showed that linkage disequilibria may be large, but did not come to a general expression for the covariance between inbred individuals when accounting for this. The close agreement between effective population sizes with and without accounting for linkage disequilibria due to selection (Table 2) suggests that the effects of linkage disequilibria due to natural selection may be small. Further it is assumed that the rate of inbreeding is small, such that the inbreeding of the parent F_t is approximately equal to that of the offspring and to the coefficient of kinship φ_t.

In the base generation the genotypic effects can be decomposed into average effects and dominance deviations (Falconer 1981). (This decomposition of genotypic values is different from the values a and d used previously in the text but is more relevant to this particular problem.) From these V_A and V_D, the additive and the dominance variance respectively, can be defined; further define V'_D as the variance of dominance deviations for homozygotes; C'_{AD} as the covariance of the total average effect and dominance deviations of homozygotes; and D' the mean dominance deviation of homozygotes.

From the work of Gillois (1964) and Harris (1964) a general form for the covariance of two individuals i and j with genotypes G_i and G_j, descended from the same base generation, is given by:

$$\mathrm{Cov}\,(G_i,G_j)=2\varphi\,(i,j)\,V_A+2Q_3\,(i,j)\,C'_{AD}+Q_4\,(i,j)\,V'_D+H(i,j)V_D+I(i,j)D'^2,$$

where φ (i,j)=coefficient of kinship of i and j; Q_n(i,j)=probability that a random sample of n of the four genes from i and j are identical; H(i,j)=probability that neither i nor j are inbred but that i and j have two pairs of identical genes; I(i,j)=joint probability of i and j both being inbred minus the product of F(i) and F(j) [where F(i)=inbreeding coefficient of i].

If i is a parent of j, two of the four genes are obliged to be identical since a copy of one gene from i is passed to j. For simplicity these two genes will be termed 'directly identical'. Then, $\varphi(i,j)=1/4(1+3F_t)$ since the probability of sampling the directly identical genes is $\frac{1}{4}$ and for all three of the remaining samplings the probability of identity by descent is approximately F_t, where F_t=average inbreeding coefficient in generation t. $Q_3(i,j)=\frac{1}{2}F_t(1+F_t)$ because with probability $\frac{1}{2}$, in a sample of three genes both directly identical genes are sampled and a third is identical with a probability F_t, and because in the remaining samplings none of the three genes are directly identical and the probability of identity is approximately F_t^2. $Q_4(i,j)=F_t^2$ is straightforward. $H(i,j)=F_t(1-F_t)$ because the directly identical pair must be distinct from the other pair, which must be identical. Finally, I(i,j)=0 since the joint probability of being inbred equals approximately the product of the inbreeding coefficients F_t^2.

Therefore, the total covariance between parent and offspring is given by

$$\mathrm{Cov}_t(G_i,G_j)=\tfrac{1}{2}(1+3F_t)V_A+F_t(1-F_t)V_D+F_t^2V'_D+F_t(1+F_t)C'_{AD}.$$

This is the covariance between G_i and G_j with respect to the non-inbred base population and includes the covariance between any two individuals belonging to the same line. However, in livestock breeding the population under consideration is a single line. Therefore, the covariance of two randomly drawn individuals from the same line must be subtracted, i.e., $\mathrm{Cov}_t(G_i,G_k)$ where i and k have no specified relationship. In this case $\varphi(i,k)=F_t$, $Q_3(i,k)=F_t^2$, $Q_4(i,k)=F_t^3$, $H(i,k)=2F_t^2(1-F_t)$, and I(i,k)=0, and so

$$\mathrm{Cov}_t(G_i,G_k)=2F_tV_A+2F_t^2(1-F_t)V_D+F_t^3V'_D+2F_t^2C'_{AD}.$$

Table A1 The dominance variance without (V_D) and with complete inbreeding (V'_D), and the covariance between additive and dominant effects with complete inbreeding C'_{AD}. The depression with complete inbreeding is 1

No. loci	V_D	V'_D	Coefficient of variation			
			0.20		0.40	
			h^2=0.05	h^2=0.10	h^2=0.05	h^2=0.10
Additive genetic variance V_A:			0.002	0.004	0.008	0.016
			C'_{AD}:			
40	0.025	0.056	0.015	0.021	0.030	0.043
160	0.006	0.014	0.008	0.011	0.015	0.021

Therefore, the covariance of parent i with offspring j within lines is given by

$$Cov_{w_i}(G_i,G_j)=(1-F_i)[\tfrac{1}{2}V_A+F_{i(1-2F_i)}V_D+F_iC'_{AD}]. \quad (A1)$$

All coefficients of non-additive terms contain F_i, hence, their contributions are small if inbreeding levels are low.

Values of V_D, V'_D and C'_{AD} for traits with small heritability and high coefficients of variation and inbreeding depression, as is the case for fitness traits, are shown in Table A1. Because $V_D<V'_D$ in Table A1, the sum of the second and third term of Eq. (A1) is positive for all F_i. Further, $C'_{AD}>0$, such that $Cov_{w_i}(G_i,G_j)>\frac{1}{2}(1-F_i)V_A$, so that $\frac{1}{2}(1-F_i)V_A$ may be used as a conservative underestimate of the covariance between parent and offspring.

Appendix 2

Reduction in variance due to natural selection

Artificial selection usually selects certain animals as parents and rejects others. Reduction in variance is then obtained by calculating variances of the selected parents. With natural selection, virtually all parents have offspring, but the numbers of offspring differ. Therefore, reduction in variance due to natural selection is obtained by weighting the parental values of the trait by the number of offspring. The relative number of offspring of selected parent i is w_i, which has mean $\bar{w}$ and variance V_w. It is assumed here that selection for production preceded natural selection, but the derivation also holds if there was no selection for production. Let X denote a trait after selection for production and X_i^* be the deviation of trait X from its mean, i.e., $E(X_i^*)=0$. Now, X_i^* is decomposed as:

$$X_i^*=b(w_i-\bar{w})+R_i,$$

where $b=Cov(w_i,X_i^*)/V_w$. It is assumed that R_i does not depend on w_i, which will approximately hold for distributions close to normal. Hence, at least approximately, $E(R_i)=0$ and $Var(R_i)=(1-r^2)V_X$, where r is the correlation between w_i and X_i^* and $V_X=Var(X_i^*)$. The variance of X_{i*} when weighted by the number of offspring, i.e., accounting for natural selection, is:

$$\begin{aligned}Var(X_i^{**})&=\sum w_iX_i^{*2}/(\bar{w}N)-[\sum w_iX_i^*/(\bar{w}N)]^2\\&=\sum w_i(w_i-\bar{w})^2b^2/(\bar{w}N)+Var(R_i)-[\sum w_i(w_i-\bar{w})b/(\bar{w}N)]^2\\&=\sum(w_i-\bar{w})^3b^2/(\bar{w}N)+\sum(w_i-\bar{w})^2b^2/N+Var(R_i)\\&\quad-[\sum(w_i-\bar{w})^2b/(\bar{w}N)+(\sum(w_i-\bar{w})b/N)]^2.\end{aligned}$$

The first term will be approximately zero if the distribution of w_i is approximately symmetric, i.e., the third central moment is approximately zero. In the last term $\sum(w_i-\bar{w})=0$, because the sum of deviations from the mean is zero. Hence approximately,

$$\begin{aligned}Var(X_i^{**})&=b^2V_w+Var(R_i)-b^2V_w^2/\bar{w}^2\\&=r^2V_X+(1-r^2)V_X-Cov^2(w_i,X_i^*)/\bar{w}^2\\&=V_X-Cov^2(w_i,X_i^*)k/V_w,\end{aligned}$$

where $k=V_w/\bar{w}^2$. Note that X_i can be w_i and the result will still hold with variances replacing covariances.

References

Abplanalp H, Peterson SJ, Okamoto S, Napolitano D (1984) Heterosis, recombination effects and genetic variation of egg composition in inbred lines of White Leghorns and their crosses. Br Poultry Sci 25:361–367

Beilharz RG (1982) The effect of inbreeding on reproduction in mice. Anim Prod 34:49–54

Bulmer MG (1971) The effect of selection on genetic variability. The Am Nat 105:201–211

Cunningham CP, Moen RA, Gjedren T (1970) Restriction of selection indices. Biometrics 26:67–70

heterosis of White Leghorns strain crosses. Can J Anim Sc 67:663–680

Falconer DS (1981) Introduction to quantitative genetics. Oliver an Boyd, Longman, London

Fisher RA (1929) The genetical theory of natural selection. Dove Publications, New York

Frankham R (1990) Are responses to artificial selection for repro ductive fitness characters consistently asymmetrical? Genet R 56:35–42

Frankham R, Yoo BH, Sheldon BL (1988) Reproductive fitness ar artificial selection in animal breeding: culling on fitness preven a decline in reproductive fitness in lines of *Drosophila melan gaster* selected for increased inebriation time. Theor Appl Ge et 76:909–914

Franklin IR (1980) Evolutionary change in small populations. I Soulé ME, Wilcox BA (eds) Conservation biology: an evolutio ary-ecological perspective. Sinauer Associates, Sunderlan Mass., 135–150

Gillois M (1964) La relation d'identité en génétique. Thesis, Un versity of Paris

Gibson JP, Kennedy BW (1990) The use of constrained selection i dexes in breeding for economic merit. Theor Appl Gen 80:801–805

Goddard ME (1983) Selection indices for non-linear profit functio Theor Appl Genet 64:339–344

Goddard ME (1992) Optimal effective population size for the gl bal population of black and white dairy cattle. J Dairy S 75:2902–2911

Goddard ME, Smith C (1990) Optimum number of bull sires in dai cattle breeding. J Dairy Sci 73:1113–1122

Hill WG (1982) Rates of change of in quantitative traits fro fixation of new mutations. Proc Natl Acad Sci USA 79:14 145

Harris DL (1964) Genotypic covariance between inbred relativ Genetics 50:1319–1348

Henderson CR (1984) Applications of linear models in animal bre ing. University of Guelph, Guelph, Canada

Jacquard A (1972) The genetic structure of populations. Sprin Verlag, Berlin

Latter BDH, Robertson A (1962) The effects of inbreeding a artificial selection on reproductive fitness. Genet Res 3:11 218

Lynch M, Gabriel W (1990) Mutation load and the survival of sm populations. Evolution 44:1725–1737

Meuwissen THE (1989) A deterministic model for the optimizat of dairy cattle breeding. Anim Prod 49:193–202

MacNeil MD, Dearborn DD, Cundiff LV, Dinkel CA, Gregory (1989) Effects of inbreeding and heterosis in Hereford fema on fertility, calf survival and preweaning growth. J Anim 67:895–901

Narain P, Robertson A (1969) Limits and duration of response to lection in finite populations: the use of transition probability trices. Indian J Hered 1:1–19

Nicholas FW, Smith C (1983) Increased rates of genetic chang dairy cattle by embryo transfer and splitting. Anim P 36:341–353

Pearson K (1903) Mathematical contributions to the theory of e lution. XI. On the influence of natural selection on the varia ity of and correlation of organs. Phil Trans Roy Soc L A200:1–66

Quinton M, Smith C, Goddard ME (1992) Comparison of selec methods at the same level of inbreeding. J Anim 70:1060–1067

Robertson A (1966) A mathematical model of the culling proces dairy cattle. Anim Prod 8:95–108

Smith C (1984) Rates of genetic change in farm livestock. Res Agric 1:79–85

Soulé IR (1980) Thresholds for survival: maintaining fitness and olutionary potential. In: Soulé ME, Wilcox BA (eds) Conse tion biology: an evolutionary-ecological perspective. Sinauer sociates, Sunderland, Mass., 151–170

Weir BS, Cockerham CC (1974) Behavior of pairs of loci in finite monoecious populations. Theor Pop Biol 6:323–354
Wiener G, Lee GJ, Woolliams JA (1992a) Effects of rapid inbreeding and of crossing of inbred lines on the body weight growth of sheep. Anim Prod 55:89–99
Wiener G, Lee GJ, Woolliams JA (1992b) Effects of rapid inbreeding and of crossing of inbred lines on the growth of linear body dimensions of sheep. Anim Prod 55:101–114
Wiener G, Lee GJ, Woolliams JA (1992c) Effects of rapid inbreeding and of crossing of inbred lines on conception rate, prolificacy and ewe survival in sheep. Anim Prod 55:115–121
Woodard AE, Abplanalp H, Snyder L (1982) Inbreeding depression in the red-legged Partridge. Poultry Sci 61:1579–1584
Wray NR, Thompson R (1990) Prediction of rates of inbreeding in selected populations. Genet Res 55:41–54
Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Paper 29

J. Anim. Breed. Genet. 115 (1998), 39–51
© 1998 Blackwell Wissenschafts-Verlag, Berlin
ISSN 0931–2668

Ms. received: 06.03.1997

[1]Scottish Agricultural College, West Mains Road, Edinburgh and [2]Roslin Institute (Edinburgh), Roslin, Midlothian, Scotland, UK

The use of Mendelian indices to reduce the rate of inbreeding in selection programmes

By B. GRUNDY[1], Z. W. LUO[2], B. VILLANUEVA[1] and J. A. WOOLLIAMS[2]

Introduction

The design of an optimal breeding programme with selection over several generations has to balance the trade-off between the short term gain to be achieved and the longer term gain that would subsequently be possible. The conflict between short and long term gains occurs because genetic improvement in selection schemes is usually associated with an increased rate of inbreeding (ROBERTSON 1961; DEMPFLE 1975) and therefore with a decreased genetic variance available for further improvement.

Best Linear Unbiased Prediction (BLUP) evaluation procedures lead to a higher rate of inbreeding and a larger magnitude of decrease in genetic variance than less accurate procedures (BELONSKY and KENNEDY 1988; QUINTON et al. 1992; VERRIER et al. 1993). Even greater increases in the rate of inbreeding are expected in selection schemes using techniques which increase reproductive capacity such as MOET (Multiple Ovulation and Embryo Transfer) (LAND and HILL 1975; NICHOLAS and SMITH 1983).

Many recent studies have been focused on developing theories of the prediction of inbreeding in selected populations (WRAY and THOMPSON 1990; WOOLLIAMS et al. 1993; WRAY et al. 1994; SANTIAGO and CABALLERO 1995) and on investigating strategies based on altering the weight given to family information in the selection criteria for controlling rates of inbreeding in selection programmes (TORO and PEREZ-ENCISO 1990; GRUNDY and HILL 1993; VERRIER et al. 1993; WRAY and GODDARD 1994; VILLANUEVA et al. 1994; BRISBANE and GIBSON 1995; LUO et al. 1995; MEUWISSEN 1997). However, these studies have not explicitly considered the optimum weight given to family information in the evaluation procedure for restricting inbreeding with minimal losses in response. VILLANUEVA and WOOLLIAMS (1997) have investigated optimum index weights when the selection criterion includes information on the candidate and its sibs.

WOOLLIAMS and THOMPSON (1994) have shown that the breeding value of an individual can be expressed in terms of the Mendelian components of itself and its ancestors. By altering the weight given to the ancestral components the amount of family information considered in the selection criterion can be varied and hence the rates of inbreeding and response balanced. They indicated that the implicit weights in BLUP evaluation were $1/2^N$ where N is the number of generations separating candidate and ancestral generation. They also described various means of altering these weights.

In the present paper, rates of response and inbreeding from selection on different selection indices based on the animal model BLUP are investigated via stochastic simulation. The indices differ in the weights given to Mendelian sampling terms of individuals born at different generations (and they are designed for reducing rates of inbreeding and consequently loss in genetic variance). It is shown how family information can be utilized in selection schemes so that short- and medium-term genetic response can be maintained while long-term genetic response may be improved. Selection procedures previously proposed

for controlling inbreeding that alter the weighting of family information (GRUNDY and HILL 1993; VERRIER et al. 1993) are discussed in the same context as Mendelian indices.

Methods

Stochastic computer simulations were used to model populations with discrete or overlapping generations. A single trait determined by an additive infinitesimal model with a heritability of 0.35 was considered. The trait was recorded in both sexes. True breeding values of the individuals in the base population (nine males and 18 females) were obtained from a normal distribution with mean zero and variance σ^2_A. Phenotypic values were obtained by adding a normally distributed environmental component with mean zero and variance $(1-\sigma^2_A)$. True breeding values for animals from subsequent years were generated as one-half the sum of the true breeding values of the animal's sire and dam plus a Mendelian sampling term taken from a normal distribution with mean zero and variance σ^2_M, where $\sigma^2_M = (1/2)(1-(F_s+F_d)/2)\sigma^2_A$ and F_s and F_d are the inbreeding coefficients of the sire and the dam, respectively. Inbreeding coefficients were obtained using the algorithm proposed by MEUWISSEN and LUO (1992). At each generation, animals with the highest index values (nine males and 18 females) were selected and mated at random under a hierarchical design. Each sire was mated to two dams. Alternative Mendelian indices were used as selection criteria and they are described later.

At each generation (with discrete generations) or each year (with overlapping generations), and for each replicate, the mean and the variance of the true breeding values and the average coefficient of inbreeding were calculated for the animals born in that generation or year. In the overlapping generation model the generation intervals for males and females were calculated as the average age of the parents when the offspring were born.

Schemes with discrete generations

A population of a constant size of 108 animals with equal frequency of each sex was simulated. Each mating pair produced three male and three female offspring. Each scheme was run for 25 generations (25 years) and replicated 250 times.

Schemes with overlapping generations

The overlapping generation model simulated a beef MOET nucleus scheme which was the same as that investigated in VILLANUEVA et al. (1994). The trait under selection was recorded in both sexes at the end of test at around 400 days of age. Unrelated animals of 2-, 3- and 4-years-of-age were generated to constitute the base population. Genetic evaluation was carried out twice every year (6 months was defined as an evaluation period). In each evaluation period, donors were flushed three times with a 2-month time interval between consecutive flushes. The unselected cows were culled from the herd. The number of transferable embryos collected from an individual flush was generated from a negative binomial distribution with mean, coefficient of variation and repeatability of 5.0, 1.25 and 0.23, respectively. These values were obtained from analyses of extensive data on dairy and beef embryo recovery experiments (see WOOLLIAMS et al. 1995 and VILLANUEVA et al. 1994 for more details). Embryos survived until calving with a fixed probability of 0.55 following a survey study (LUO et al. 1994). Each simulation was run for 25 years and was replicated 100 times.

Mendelian indices

In all selection schemes reported here the genetic evaluation of animals was performed using an animal model BLUP with the population mean being the only fixed effect considered.

Estimates of the Mendelian terms were calculated as the difference between individual estimated breeding value and average parental breeding value derived from the BLUP procedure. As shown by WOOLLIAMS and THOMPSON (1994) by systematically expressing each breeding value encountered as one descends a pedigreea a breeding value can be expressed solely in terms of Mendelian components, such that an estimated breeding value of ith individual (a_i) obtained within a BLUP framework can be decomposed into the estimates of the Mendelian components of itself and its ancestors as

$$a_i = m_i + \sum_{t=1}^{T} \sum_{j=1}^{2^t} m_{ij,t} c^t \quad (1)$$

where m_i is the estimated Mendelian sampling term for the ith individual, $m_{ij,t}$ is the estimated Mendelian sampling term for the jth ancestor of the ith individual with t generations separating individual and ancestor. With BLUP, c is equal to 0.5 and T (the number of generations to be included in the index) is equal to N, the number of generations of selection. It must be noted that the Mendelian sampling term estimates are identical to the estimated breeding values for the animals in the base population.

The decomposition of the estimated breeding value in the form of (1) illustrates how BLUP selection relies on the information contributed from pedigree. Moreover, altering the weights given to Mendelian terms of ancestors of different generations does offer a flexible approach to achieve the optimum index for selection from which rates of inbreeding may be reduced without causing substantial decrease in genetic gain. In the present simulation study two options for altering the Mendelian contributions were considered through changing values of c and T.

In order to reduce the family information accumulated in the index, a truncation approach may be taken by which information after a given generation is excluded. This is the form of the Truncated Mendelian Index (TM) which can be expressed as (1) where T is less than N. Schemes where T takes the value of 1, $TM(1)$ (i.e. only terms from the individual and parental generation are included) or 2, $TM(2)$ (i.e. only individual, parental and grandparental terms are included) are investigated.

The breeding value can be expressed as a weighted sum of Mendelian terms that can be expressed as a geometric series in c^t. When c is less than unity, these weights incrementally decrease each generation and thus reduce the weighting given to family information from increasingly distant generations. By allowing c to take values less than a half, pedigree information can be progressively discounted at a greater rate than with BLUP. Hence this family of indices are termed Geometric Mendelian Index (GM) and have the form of (1) with c taking values less than 0.5 and $T = N$. Schemes with c taking values of 0, 0.1, 0.2, 0.3 and 0.4 are investigated. When c is zero, selection is based on a candidate's Mendelian term estimate only. Under the discrete generation case, this is a form of within family selection.

A modification of the Mendelian indices for use with overlapping generations

In standard BLUP selection the total summed weight (γ) of ancestral Mendelian components to a breeding value increases each generation by a constant value of one assuming no reduction in variance due to inbreeding. Therefore the later generations will have a reduced component compared to earlier generations. Assuming a constant rate of progress per generation, then γ on average accounts for a proportion of $(N-1)/N$ of an individual breeding value at generation N. For example at the base generation, no ancestral information is contributed to the breeding value, and by generations two, three and four γ accounts on average for one-half, two-thirds and three-quarters of the breeding value, respectively.

For the TM schemes the ratio of γ to $(N-1)/N$ is progressively reduced each generation. As a consequence of that, breeding values from different generations are no longer strictly comparable, as they are with BLUP selection. Breeding values from individuals born in

later generations will have been subject to larger reductions than the breeding values of their older counterparts. In a population with discrete generations this problem does not arise as selection choice is made within a single cohort.

It is desirable, therefore, when dealing with overlapping generations to correct for this loss in progress to date, but to do so without increasing the rate of inbreeding. One way of addressing this imbalance is to replace the individual reduction in breeding value with a correction factor derived from a more general source weighted by the average degree of loss incurred. The average loss, however, is dependent on the index used, so that specific corrections are needed.

The correction factor used for $GM(c)$ was derived as follows. Assuming an average generation interval the average loss in the ancestral component of the breeding value each year can be expressed as

$$\gamma = \sum_{m=1}^{N} 2^m c^m$$

where m and N are respectively the average number of years divided by the generation interval separating individual and ancestral generation and base generation. The average generation interval can be obtained from the standard BLUP evaluation. For $TM(T)$ schemes γ is simply $(N-T)/N$. In addition, the average ancestral component, $\bar{x}$, is calculated as the average breeding value minus the average Mendelian deviation. Hence, the correction factor to be added in both cases is $(1-\gamma)\bar{x}$. With this correction the degree of progress to date discarded is accounted for in an appropriate manner. Comparisons of schemes with and without a correction factor are undertaken for overlapping generation schemes.

Alternative indices

The lower the heritability of the quantitative trait under selection, the more important the family information is in the BLUP evaluation and thus the higher level of inbreeding expected. GRUNDY and HILL (1993) suggested the use of a heritability which is artificially biased upwards in the BLUP evaluation for restricting family information on which the BLUP selection was relied. In the present study, selection schemes with the biased BLUP evaluation were denoted by BH. Additive genetic variances artificially inflated in BLUP evaluation by 2 ($BH(2)$) and 3 ($BH(3)$) times were simulated.

VERRIER et al. (1993) suggested another form of selection index reconstructed from BLUP evaluation in which the emphasis given to the parental (and consequently the ancestral) information was reduced. The modified index, denoted by VCF in the present study, was of the form $CF_i = m_i + (a_s + a_d)/2 - \lambda(a_s + a_d)$ where a_s and a_d are the estimated breeding values of the sire and dam obtained from BLUP, and λ takes values between 0 and 0.5. Note that when λ takes the value of zero the selection is on standard BLUP, whereas when λ takes the value of one-half, selection is based on within family deviations. In the simulations, λ takes the value of a quarter. This index can be expressed in Mendelian terms by setting $c^t = (1/2 - \lambda)2^{-t+1}$ in equation (1).

Results

Schemes with discrete generations

Figure 1 shows the relative cumulative genetic gain over the period of selection observed when using the different index selection schemes compared with standard BLUP selection (i.e. the graphs show response by index selection divided by response by BLUP times 100). Each line on the graph can be divided into three sections: an initial period where the modified indices behave the same as standard BLUP; a period when the index schemes

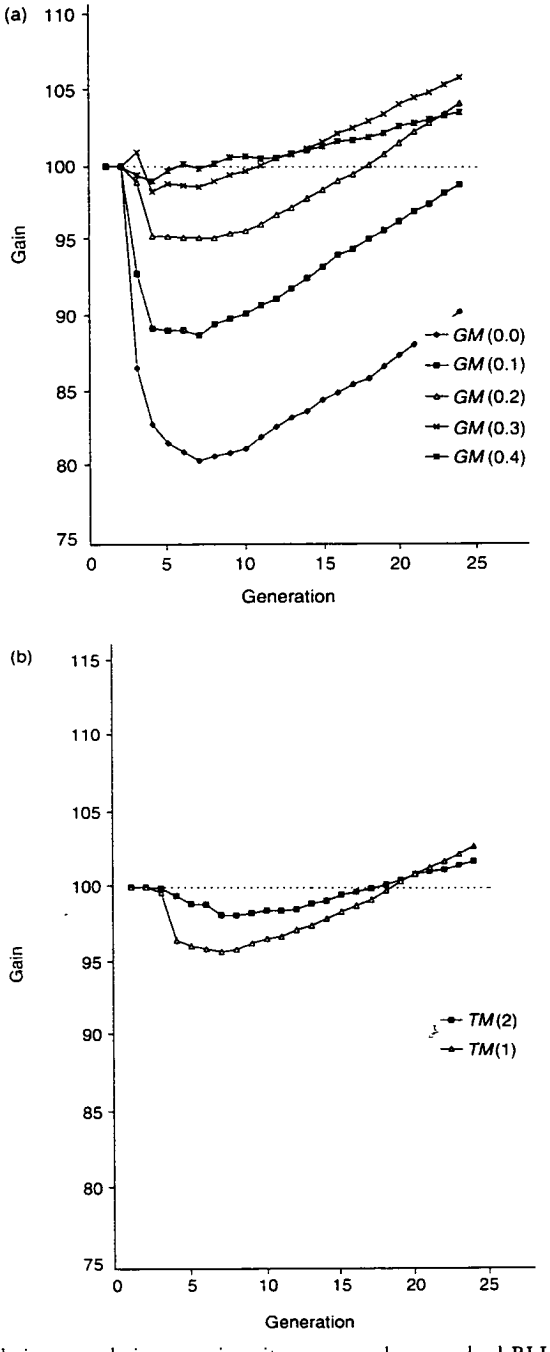


Fig. 1. The average relative cumulative genetic gains compared to standard BLUP evaluation for the simulated breeding schemes with discrete generations: (A) the geometric Mendelian index (*GM*) with the weight *c* varying from 0.1 to 0.4, (B) the Mendelian term index (*TM*) including information on one or two ancestral generations (C) the selection index suggested by VERRIER et al. (*VCF*) and the index using biased heritability estimates (*BH*).

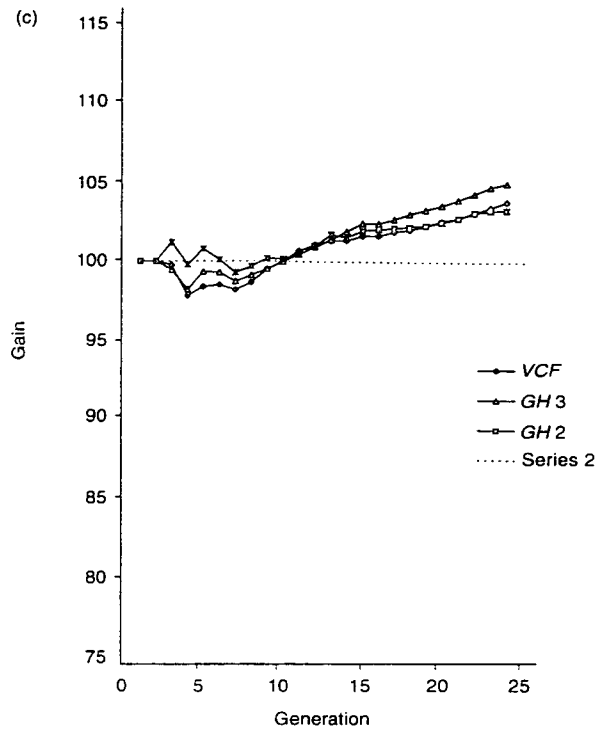


Fig. 1.—continued.

show a lower rate of response; and a period when the index schemes show an increased rate of response.

As expected the genetic responses to the first generation of selection using all the modified indices were identical to the response from standard BLUP since in the base population there is no family information. Thereafter, compared to the standard BLUP selection, a sharp decrease in the response was observed in the following generations of selection using the various indices. The size of initial reduction in the rate of response was proportional to the decrease in the amount of family information included in the selection index.

It is worth noting that these first two stages are an artefact of the selection being commenced when there was no family information available. If the selection on the modified index would have been introduced later on the relative gain of schemes using indices would have started at the lowest point on the graph. In the later generations although the rate of genetic response in all selection schemes has gradually decreased over time, the relative rate of response in the index schemes has decreased at a slower rate than BLUP.

For the *GM* indices the smaller the difference of c from 0.5, the greater the reliance upon the family information and thus the smaller the initial reduction in the genetic response (Fig. 1a). After 25 generations compared to the standard BLUP selection, the cumulative genetic gain increased by 6.1, 7.5 and 4.9 % for the *GM* schemes with c equal to 0.2, 0.3 and 0.4, respectively. Moreover, *GM* schemes that show greater rates of responses in later generation are those with a proportionally greater reduction in family information. The rate of response averaged across generations 15 to 25 are shown in Table 1. Differences in the rates of response for the various schemes were highly significant.

The genetic response in schemes with selection on *TM* indices was maintained at the same

Table 1. **Genetic progress (in units of the initial phenotypic standard deviation), inbreeding (%), and additive genetic variance (σ^2_{A25}), their standard errors observed after 25 generations of selection in the simulated breeding schemes with discrete generations. The selection methods performed in these schemes were: the standard BLUP evaluation (BLUP); the Geometric Mendelian Index (*GM*) with the weight *c* varying from 0.1 to 0.4; the Truncated Mendelian Index (*TM*) including one or two generations of ancestral information; the selection index suggested by VERRIER et al. (*VCF*) and the index proposed by GRUNDY and HILL (*BH*)**

Selection method	Genetic progress		Inbreeding		
	G_{25}	ΔG_{15-25}	F_{25}	ΔF_{15-25}	σ^2_{A25}
BLUP	7.18 ± 0.03	0.125 ± 0.002	68.46 ± 0.32	4.98 ± 0.24	0.095 ± 0.002
GM(0.0)	6.59 ± 0.03	0.220 ± 0.003	32.53 ± 0.10	1.65 ± 0.14	0.214 ± 0.003
GM(0.1)	7.24 ± 0.03	0.203 ± 0.003	37.35 ± 0.13	1.98 ± 0.16	0.188 ± 0.002
GM(0.2)	7.62 ± 0.03	0.187 ± 0.003	47.23 ± 0.17	2.45 ± 0.17	0.165 ± 0.002
GM(0.3)	7.72 ± 0.03	0.164 ± 0.002	51.94 ± 0.24	3.14 ± 0.19	0.138 ± 0.002
GM(0.4)	7.53 ± 0.03	0.143 ± 0.002	60.20 ± 0.27	4.03 ± 0.22	0.117 ± 0.002
TM(1)	7.52 ± 0.03	0.169 ± 0.002	50.22 ± 0.17	2.92 ± 0.19	0.149 ± 0.002
TM(2)	7.47 ± 0.03	0.144 ± 0.002	60.59 ± 0.24	4.04 ± 0.22	0.118 ± 0.002
VCF	7.53 ± 0.03	0.149 ± 0.002	58.36 ± 0.33	3.90 ± 0.22	0.123 ± 0.002
BH(2)	7.50 ± 0.03	0.143 ± 0.002	60.23 ± 0.34	4.08 ± 0.22	0.117 ± 0.002
BH(3)	7.63 ± 0.03	0.154 ± 0.002	55.74 ± 0.37	3.53 ± 0.20	0.131 ± 0.002

level as the gain from BLUP selection in the initial two to three generations and then decreased in the following five to seven generations of selection (Fig. 1b). The magnitude of reduction in the response depended on the amount of information included in the index. After the relative cumulative genetic responses reached a minimum, they gradually increased so that by generation 25 they were highly significantly greater than the response obtained in the BLUP scheme by 4.7 and 4.0 % for the schemes with *TM*(1) and *TM*(2), respectively.

The pattern of change in the relative cumulative genetic gain for the scheme with *VCE* index over the selection process was similar to that for the scheme with *TM*(1) (Fig. 1c). As with the other indices smaller genetic gains were observed for the *BH* and *VCF* schemes compared to the reference scheme in the early period of the selection process, but by the end of the selection process, the relative genetic gain obtained in the *VCF* and *BH* selection schemes significantly exceeded that of the BLUP scheme. With the *BH* schemes the relative genetic gain at generation 25 compared to BLUP selection increased by 5.9 % with an inflated genetic variance of two-fold. With the *VCF* scheme the relative genetic gain compared to BLUP selection increased by 4.9 % (Table 1).

Rates of inbreeding and accumulated inbreeding levels after 25 generations of selection are shown in Table 1. The average rate of inbreeding in the last 10 generations was decreased relative to standard BLUP selection by 66.9 % to 19.1 % with *GM* selection with *c* increasing from zero to 0.4, by 41.4 and 18.9 % with *TM*(1) and *TM*(2), respectively, by 21.7 % with *VCF*, and by 29.1 % to 18.1 % with *BH* with additive genetic variance inflated by three and two times, respectively. All these differences were highly significant. There are clear trends for all schemes where increased reduction in family information lead to reduced inbreeding.

The average additive genetic variance after 25 generations of selection is shown in Table 1. In all the index schemes the genetic variance was higher than that from standard BLUP selection as the relative dependence of the selection indices on parental contributions was lower. The proportions of an increased additive genetic variance after 25 generations of selection, compared to the standard BLUP selection scheme ranged from 23 % to 125 %

Table 2. **Genetic progress (in units of the initial phenotypic standard deviation), inbreeding (%), additive genetic variance (σ^2_{A25}), and average generation interval over both sexes (L) together with their standard errors observed after 25 years of selection in the simulated beef MOET schemes. The selection methods performed in these schemes were: the standard BLUP evaluation (BLUP); the Geometric Mendelian Index (*GM*) with the weight c varying from 0.1 to 0.4; the Truncated Mendelian Index (*TM*) including one or two generations of ancestral information, the selection index suggested by VERRIER et al. (*VCF*) and the index proposed by GRUNDY and HILL (*BH*)**

Selection methods	Genetic progress		Inbreeding			
	G_{25}	ΔG_{15-25}	F_{25}	ΔF_{15-25}	σ^2_{A25}	L
BLUP	4.69 ± 0.03	0.203 ± 0.002	41.73 ± 0.87	2.27 ± 0.12	0.246 ± 0.004	3.07 ± 0.03
GM(0.0)	2.41 ± 0.04	0.087 ± 0.003	21.99 ± 0.65	1.12 ± 0.06	0.290 ± 0.005	7.47 ± 0.15
GM(0.1)	2.62 ± 0.04	0.086 ± 0.002	22.36 ± 0.64	1.03 ± 0.06	0.286 ± 0.006	7.58 ± 0.15
GM(0.2)	2.98 ± 0.04	0.095 ± 0.003	25.73 ± 0.67	1.13 ± 0.08	0.280 ± 0.005	7.29 ± 0.16
GM(0.3)	3.41 ± 0.04	0.111 ± 0.004	30.05 ± 0.66	1.44 ± 0.08	0.281 ± 0.006	6.61 ± 0.15
GM(0.4)	4.05 ± 0.03	0.146 ± 0.003	36.72 ± 0.76	1.81 ± 0.09	0.256 ± 0.005	4.85 ± 0.10
TM(1)	2.91 ± 0.04	0.104 ± 0.003	29.13 ± 0.66	1.39 ± 0.09	0.293 ± 0.006	6.89 ± 0.16
TM(2)	3.38 ± 0.05	0.114 ± 0.003	33.76 ± 0.77	1.53 ± 0.09	0.273 ± 0.005	6.29 ± 0.15
VCF	4.25 ± 0.04	0.177 ± 0.003	35.30 ± 0.72	1.94 ± 0.09	0.247 ± 0.004	3.95 ± 0.08
BH(2)	4.61 ± 0.04	0.203 ± 0.003	32.29 ± 0.78	1.73 ± 0.09	0.248 ± 0.004	3.20 ± 0.03
BH(3)	4.55 ± 0.04	0.198 ± 0.003	27.14 ± 0.64	1.41 ± 0.07	0.250 ± 0.004	3.38 ± 0.03

Table 3. **Genetic progress (in units of the initial phenotypic standard deviation), inbreeding (%), additive genetic variance (σ^2_{A25}), and average generation interval over both sexes (L) together with their standard errors observed after 25 years of selection in the simulated beef MOET schemes. The selection methods performed in these schemes were: the standard BLUP evaluation; the Geometric Mendelian Index (*GM*) with correction* and the Truncated Mendelian Index (*TM*) with correction***

Selection methods	Genetic progress		Inbreeding			
	G_{25}	ΔG_{15-25}	F_{25}	ΔF_{15-25}	σ^2_{A25}	L_{m+f}
BLUP	4.69 ± 0.03	0.203 ± 0.002	41.73 ± 0.87	2.27 ± 0.12	0.246 ± 0.004	3.07 ± 0.03
GM(0.0)*	3.82 ± 0.04	0.180 ± 0.002	29.30 ± 0.64	1.58 ± 0.08	0.246 ± 0.004	3.17 ± 0.03
GM(0.1)*	4.15 ± 0.03	0.187 ± 0.003	29.00 ± 0.53	1.62 ± 0.07	0.251 ± 0.005	3.12 ± 0.03
GM(0.2)*	4.28 ± 0.04	0.181 ± 0.002	32.21 ± 0.64	1.68 ± 0.07	0.243 ± 0.005	3.11 ± 0.03
GM(0.3)*	4.34 ± 0.03	0.184 ± 0.003	34.23 ± 0.67	1.75 ± 0.08	0.244 ± 0.004	3.04 ± 0.03
GM(0.4)*	4.54 ± 0.03	0.193 ± 0.003	37.95 ± 0.81	1.90 ± 0.10	0.247 ± 0.005	3.00 ± 0.03
TM(1)*	4.14 ± 0.04	0.181 ± 0.003	37.47 ± 0.81	1.91 ± 0.10	0.247 ± 0.004	3.16 ± 0.03
TM(2)*	4.27 ± 0.04	0.174 ± 0.003	38.24 ± 0.74	2.03 ± 0.10	0.252 ± 0.004	3.18 ± 0.03

for the *GM* schemes with c decreasing from 0.4 to zero; and were 22 and 57 % for the *TM* selection schemes and 29, 23 and 38 % for the *VCF*, *BH*(2) and *BH*(3) selection schemes, respectively.

Schemes with overlapping generations

The cumulative genetic gains achieved with overlapping generation schemes are shown in Tables 2 and 3 for the different selection methods. Significantly lower rates of genetic gain

were observed for the uncorrected Mendelian schemes relative to BLUP selection. These lower rates of gain were associated with increased generation intervals (Table 2).

With *GM*(0.1) the generation interval was higher than that obtained with BLUP by an extra of 4.5 years. This trend of increase in generation interval and decrease in response was consistent across all schemes where the ancestral component was reduced in the selection criteria. In contrast, schemes in which the correction was used, although showing the same overall trends, were more competitive showing rates of response more similar to those from BLUP selection. The average cumulative progress, from years 15 to 25, as a percentage of those obtained by BLUP selection ranged from 89 to 95 % for *GM* schemes with *c* varying from zero to 0.5, whereas, the equivalent figures from the schemes with no correction applied ranged from 43 to 72 %.

With the uncorrected *TM*(1) and *TM*(2) schemes the generation intervals were respectively 124 and 104 % greater than that obtained with BLUP, an extra of 3.82 years for the *TM*(1) scheme. In the corrected *TM* schemes, generation interval was decreased to a similar level to that obtained with standard BLUP. Cumulative responses after 25 years, as a percentage of those obtained by BLUP selection were 88 and 91 % with *TM*(1) and *TM*(2), respectively, whereas, the comparable figures from the schemes with no correction were 62–72 %.

No corrections were needed for the *BH* scheme which yielded higher long term gains than BLUP for both discrete and overlapping generation models. Generation intervals were increased by 4 and 10 % for *BH*(2) and *BH*(3), respectively, but these represent relatively small changes. In the *VCF* scheme, generation intervals were increased by 0.9 years. Table 2 shows that the difference in mean cumulative response between the *BH* schemes and BLUP to be less than 3 % for both *BH*(2) and *BH*(3). The *VCF* scheme gave a cumulative response 9.4 % lower than BLUP selection.

The rate of inbreeding over the last 10 generations of selection and the average inbreeding coefficient after 25 generations of selection is shown in Tables 2 and 3 for the different selection methods used in the MOET schemes. The final inbreeding level of the standard BLUP selection scheme was reduced by 12 % to 46 % with *GM* selection with *c* decreasing from 0.4 to 0.1; by 31 % or 12 % with the *TM*(1) or *TM*(2) selection, respectively and by 23–35 % with the *VCF* or *BH* selection indices. Although there was a large variation in the inbreeding reduction amongst the different selection indices, these differences were all highly significant. The magnitude of the reduction in the rate of inbreeding compared to the standard BLUP selection was inversely proportional to the amount of family information included in the modified selection indices.

The patterns in the dynamics of genetic variance were similar to those described for discrete generations. After 25 years (around eight generations) of selection in the schemes with overlapping generations, the increased proportions of additive genetic variance compared to that maintained in the BLUP selection scheme, ranged from 4 to 18 % in the *GM* schemes and from 11 to 19 % in the *TM* schemes. There were no significant differences in genetic variance for the *VCF*, *BH* and corrected *GM* and *TM* selection schemes (Table 3). The magnitude of the genetic variance maintained at the end of selection process was dependent on the parental contributions to the selection index.

Discussion

By partitioning the family information within an index both short and longer term rates of inbreeding can be altered. An advantage in the short term gain does not necessarily imply greater long term gains. In the present study the use of two modified selection indices, the *TM* index and the *GM* index, which are derived from the BLUP evaluation have been investigated. In principle, the use of these indices enables the family information contained in the standard BLUP to be used in a flexible manner, such that, ideally, only the minimum amount of family information is included in the selection criteria in order to sustain the

short term genetic response, and to increase the long term response in comparison to standard BLUP selection by reducing the rate of inbreeding.

Genetic response is influenced by several factors including the intensity of selection, the available genetic variance and the accuracy of the selection index. These factors change at different rates with different selection schemes and ultimately influence the responses that are observed in both the short and longer term. By altering the amount of family information in the selection index the intensity of selection and the accuracy are altered. The applied selection differential is greater with the modified index methods than with BLUP because the use of family information increases the intraclass correlation and increases the correlations in the estimated breeding values (EBVs) leading, in finite populations, to a reduced selection differential (HILL 1976; RAWLINGS 1976; MEUWISSEN 1991). With modified indices the correlation between index values is reduced compared to BLUP. This effect is likely to be small but related to the amount of family information that is discarded. The accuracy of selection is lower when using modified indices.

In early generations the variance will be similar for all schemes but it will obviously be reduced more with the most efficient methods. Again for the more extreme cases this effect may be quite large. However, the advantage of a relatively higher variance in the index schemes does not compensate for the decreased accuracy of selection. In later generations, the impact of inbreeding becomes an increasing factor reducing the genetic variance still further; as the index methods show lower rates of inbreeding they maintain relatively more variance. When this second factor outweighs the disadvantages of the reduced accuracy, the index scheme becomes more efficient. This does, however, indicate the limitations of the system for relatively large populations as this point will occur later as the population size increases. Ultimately, the modified indices will be advantageous but this, especially with large populations and low family information could take longer.

The present research includes two population models: discrete generations and overlapping generations. In the former, selection candidates shared a constant number of parental Mendelian term contributions in their EBVs from BLUP evaluation, and thus the number of parental Mendelian term estimates which are available to reconstruct the modified BLUP indices (*GM*, *TM* and *VCF*) are identical amongst all selection candidates. However, in the simulated MOET schemes, selection candidates have an overlapping age distribution and thus these modified indices may contain different levels of parental contributions among the selection candidates. The consequence of this was that the generation interval was increased and the response obtained was reduced. For the schemes using Mendelian indices the benefit in the reduction in the rate of inbreeding was of the same magnitude as the loss in response which would limit their application to populations with overlapping generations. The correction for overlapping generations was shown to be effective at breaking this relationship, but this was at the cost of increased inbreeding.

In schemes with discrete generations selection was carried out for 25 generations while selection in the simulated MOET schemes (with overlapping generations) was performed for only 25 years which were equivalent to about eight generations for the standard BLUP scheme. Thus, the consequences of using the modified BLUP indices was more effectively revealed in the discrete generation schemes as the benefits of using these indices increase over time. Long-term genetic response can be substantially improved by the use of the *GM* and *TM* as first suggested in WOOLLIAMS and THOMPSON (1994) since a considerable amount of inbreeding is avoided and a higher level of additive genetic variance is maintained. The selection using the *VCF* index does also show a clear superiority to the standard BLUP scheme for its improved long-term response and reduced rates of inbreeding. This agrees with the simulation results in VERRIER et al. (1993) in which similar simulation parameters were used.

Many methods have recently been developed for controlling the rate of inbreeding in selection programmes. A general conclusion to be drawn from the literature is that selection routines designed to achieve maximum gains at a fixed level of inbreeding (WRAY and

GODDARD 1994; BRISBANE and GIBSON 1994) are the most efficient. However, these routines have so far been 'black box' in their approach as they do not explicitly calculate the optimum weights given to family information whereas, index selection methods tend to have a simpler and clearer framework.

Most of the methods proposed for controlling inbreeding have considered discrete generation models. Although it is difficult to define the reduced weighting in overlapping generations this is possible given that the biased heritability method is effective in overlapping generations. Problems are specifically based on the difficulty in the control of the sampling errors when populations overlap as several cohorts can contribute to the next generation.

The results from the discrete generation model indicates a similarity between the effect of biasing the heritability and using the *GM* index, which can be interpreted as a common methodology based on influencing selection decisions through progressively altering the weight given to more distant family information. As the *BH* index sacrifices the information progressively with t generations of selection, according to a geometric series, so the *GM* selection can be recognised as a first order approximation to the *BH* index. However, the index calculations differ: for the *BH* scheme the evaluation of breeding values from all generations is simultaneous whereas for the *GM* scheme a post BLUP correction to the parental component of the breeding values is made. As a consequence, when generations overlap the *BH* scheme is more effective at maintaining response whilst reducing inbreeding than the *GM* scheme, as any differences in mean performance between individuals with different numbers of generations of ancestors are not properly accounted for with *GM*. Hence, with not correction to *GM* the rates of inbreeding can be reduced but the response suffers as indices from different generations are not properly accounted for.

In conclusion the results of the present study demonstrated that the rate of inbreeding in long-term selection schemes can be largely reduced whilst long-term genetic responses can be substantially improved using modified indices reconstructed from the animal model BLUP genetic evaluation.

Acknowledgements

This work was funded by the BBSRC and the EU. SAC also receives funds from the Scottish Office, Agriculture, Environment and Fisheries Department. Work at the Roslin Institute receives funds from MAFF.

Summary

Genetic response and inbreeding can be balanced by altering the emphasis placed on ancestral information in genetic evaluation. The breeding value of an individual can be expressed as a weighted sum of Mendelian deviations of itself and all its ancestors. With standard BLUP selection the weights given to Mendelian terms are c^T, where T is the number of generations separating ancestor and individual and c is equal to 0.5. Selection indices with altered weights are investigated using a stochastic simulation and compared for genetic response and inbreeding after 25 years of selection for discrete and overlapping generation models. The trait analysed is measured in both sexes and has a heritability of 0.35. Two selection indices are studied: the Geometric Mendelian Index (*GM*) which used values of c between 0 and 0.5, and the Truncated Mendelian Index (*TM*) which uses Mendelian terms from a limited number of generations (for example when $T = 2$ the selection criterion includes individual, parental and grandparental Mendelian terms). A consequence of reducing ancestral information is that accrued progress is disregarded. This is critical for overlapping generations as sources of information vary between individuals. This is corrected by adding a weighted average of ancestral information of the individuals generation to the index. With discrete generations, *TM* with only parental information included reduced inbreeding by 24% and increased cumulative response by 6% compared to standard BLUP by generation 25. As T increases the method tends towards BLUP. With *GM*, maximum benefits over standard BLUP were obtained at c equal to 0.3 with 23% less inbreeding and 8% and more cumulative response by generation 25. The Mendelian Indices were less effective for overlapping generations. Both

types of indices were efficient methods for increasing long-term response. Moreover, comparison with other recent selection schemes highlights common methodologies.

Zusammenfassung

Verwendung Mendel'scher Indices zur Reduktion der Inzuchtrate in Selektionsprogrammen

Zuchterfolg und Inzucht kann durch entsprechende Betonung der Ahneninformation bei Zuchtwertschätzung balanziert werden. Der Zuchtwert eines Individuums kann als gewichtete Summe Mendel'scher Abweichungen von ihm und den Verwandten ausgedrüekt werden. Standard BLUP Selektion gibt Gewichte T den Mendel'schen Größen, wobei T die Generationenzahl zwischen Individuum und Vorfahre ist und $c = 0.5$. Selektionsindices mit geänderten Gewichten werden in stochastischer Simulation untersucht und hinsichtlich Erfolg und Inzucht nach 25 Jahren Selektion mit diskreten und überlappenden Generationen verglichen. Daas Merkmal hat eine Heritabilität von 0.35 in beiden Geschlechtern, Der Geometrische Mendel'scher Index (*GM*) verwendet *C* Werte zwischen 0 und 0.5, der gestutzte (*TM*) nutzt Mendel'sche Größen für begrenzte Zahl von Generationen ($T = 2$ umfaßt Individuum, Eltern und Großeltern). Eine Konsequenz reduzierter Ahneninformation ist Verzicht auf entstandenen Fortschritt, was überlappenden Generationen kritisch sein kann, da Informationsquellen zwischen Individuen schwanken. Dem wird durch Addition eines gewichteten Durchschnittes der Ahneninformmation Rechnung getragen. Bei diskreten Generationen wird mit *TM* mit nur parentaler Information die Inzucht um 24 % reduziert und kumulativer Erfolg um 6 % im Vergleich zu Standard BLUP bei 25 Generationen gesteigert. Mit zunehmendem *T* tendiert die Methode zu BLUP. Die Vorteile werden bei GM mit $c = 0.3$ maximiert, 23 % weniger Inzucht und um 8 oder mehr % höherer kumulativer Erfolg. Die Mendel'schen Indices waren bei überlappenden Generationen weniger wirksam. Beide Arten von Indices waren hinsichtlich langfristiger Zuchterfolge wirksam und Vergleich mit anderen neueren Selektionsplänen erweist Gemeinsamkeiten.

References

BELONSKY, G. M.; KENNEDY, B. W., 1988: Selection on individual phenotype and best linear unbiased prediction of breeding value in a closed swine herd. J Anim. Sci. **66:** 1124–1131.

BRISBANE, J. R.; GIBSON, J. P., 1995: Balancing selection response and the rate of inbreeding by including genetic relationships in selection decisions. Theoret. Appl. Genet. **91:** 421–431.

DEMPFLE, L., 1975: A note on increasing the limit of selection through selection within families. Genet. Res. (Camb.) **24:** 127–135.

GRUNDY, B.; HILL, W. G., 1993: A method for reducing inbreeding with best linear unbiased prediction. Anim. Prod **56:** 427. (Abst.).

HILL, W. G., 1976: Order statistics of correlated variables and implications in genetic selection programs. Biometrics **32:** 889–902.

LAND, R. B.; HILL, W. G., 1975: Effects of population size on response to short and long term selection. Anim. Prod. **21:** 1–12.

LUO, Z. W.; WOOLLIAMS, J. A.; SIMM, G., 1994: An assessment of present and future effectiveness of embryological techniques in ruminants. Ag. Biotech. News. Infor. **6:** 13N–18N.

–; –; THOMPSON, R., 1995: Strategies for controlling rates of inbreeding in dairy MOET nucleus schemes. Anim. Sci. **60:** 379–387.

MEUWISSEN, T. H. E., 1991: Reduction of selection differentials in finite populations with nested full-half sib family structure. Biometrics **47:** 195–203.

–, 1997: Maximising the response of selection with a predefined rate of inbreeding. J. Anim. Sci. **75:** 934–940.

–; LUO, Z. W., 1992: Computing inbreeding coefficients in large populations. Genet. Select. Evol. **24:** 305–311.

NICHOLAS, F. W.; SMITH, C., 1983: Increased rates of genetic changes in dairy cattle by embryo transfer and splitting. Anim. Prod **36:** 341–353.

QUINTON, M.; SMITH, C.; GODDARD, M. E., 1992: Comparison of selection methods at the same level of inbreeding. J. Anim. Sci. **70:** 1060–1067.

RAWLINGS, J. O., 1976: Order for a special class of unequally correlated multinominal variates. Biometrics **32:** 875–877.

ROBERTSON, A., 1961: Inbreeding in artificial selection programmes. Genet. Res. (Cam.) **2:** 189–194.

SANTIAGO, E.; CABALLERO, A.; 1995: Effective size of populations under selection. Genetics **139:** 1013–1030.

TORO, M.; PEREZ-ENCISO, M., 1990: Optimisation of selection response under restricted inbreeding. Genet. Select. Evol. **22:** 93–107.

VERRIER, E.; COLLEAU, J. J; FOULLEY, J. L., 1993: Long-term effects of selection based on animal model BLUP in a finite population. Theoret. Appl. Genet. **87:** 446–454.

VILLANUEVA, B.; WOOLLIAMS, J. A., 1997: Optimisation of breeding programmes under index selection and constrained inbreeding. Genet. Res. (Camb.) **69:** 145–158.
–; –; SIMM, G., 1994: Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. Genet. Select. Evol. **26:** 517–535.
–; WRAY, N. R.; THOMPSON, R., 1993: Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. Anim. Prod. **57:** 1–13.
WOOLLIAMS, J. A.; LUO, Z. W.; VILLANUEVA, B.; WADDINGTON, D.; BROADBENT, P. J.; MCKELVEY, W. A. C.; ROBINSON, J. J., 1995: Analysis of factors affecting superovulatory response in ruminants. J. Agric. Sci., Camb. **124:** 61–70.
–; THOMPSON, R., 1994: A theory of genetic contributions. In: Proc. 5th World Congress on Genetics Applied to Livestock Production, Vol. 19. Guelph. pp. 127–134.
–; WRAY, N. R.; THOMPSON, R., 1993: Prediction of long-term contributions and inbreeding in populations undergoing mass selection. Genet. Res. (Camb.) **62:** 231–242.
WRAY, N. R.; GODDARD, M. E., 1994: Increasing long term response to selection. Genet. Select. Evol. **26:** 431–451.
–; THOMPSON, R., 1990: Prediction of rates in inbreeding in selected populations. Genet. Res. **855:** 41–54.
–; WOOLLIAMS, J. A.; THOMPSON, R., 1994: Prediction of rates of inbreeding in populations undergoing index selection. Theoret. Appl. Genet. **87:** 878–892.

Address of authors: B. GRUNDY, Scottish Agricultural College, Genetics and Behavioural Sciences, Bush Estate, Penicuik, EH26 0QE, Scotland UK.

Reasons for the different conclusions can be found from considering the formula given in Woolliams *et al.* (1993) denoted by ΔF_{LH} which, if simplified by assuming that maternal and paternal half-sib correlations are equal ($\rho_m = \rho_f = \rho$), is given by:

$$\Delta F_{LH} = (8N_s)^{-1} + (8N_d)^{-1} - (8T)^{-1} + i^2\rho(1 + 3d^{-1})(8N_s)^{-1}$$
$$= (8N_s)^{-1} + (1 + d^{-1}) - (8T)^{-1} + i^2\rho(1 + 3d^{-1})(8N_s)^{-1}.$$

For a fixed N_s, the term $(8N_s)^{-1}$ $(1 + d^{-1})$ in ΔF_{LH} is minimized when d increases. However, the selection intensity also decreases as d increases and the results of Jódar and López-Fanjul (1977) show that this balance favours $d = 1$. With selection, two additional terms are included ($(8T)^{-1}$ and $i^2\rho(1 + 3d^{-1})(8N_s)^{-1}$). The term $(8T)^{-1}$ is considered fixed in this study ($T = N/2$) while the term $i^2\rho(1 + 3d^{-1})(8N_s)^{-1}$ decreases with d but increases with i^2. Gain is proportional to i and i is less sensitive to changes in d when selection pressure is strong. Therefore, with large selection intensities an increase in d may reduce the rate of inbreeding more strongly than the rate of gain. Large values of N result in large intensities. Under these circumstances, an increase in d will tend to smaller reductions in the rate of gain but more dramatic decreases in the rate of inbreeding. Thus it is reasonable that with large values of N, potentiated by large values of ρ (obtained with moderate to high values of $h^2_{(0)}$) the optimum mating ratio may shift from $d = 1$ to $d = 2$.

Severe restrictions on inbreeding with the smallest resources can lead to optimum $d = 1$. As ΔF decreases, N_s must increase to achieve the restriction and, with fixed N, i must be reduced in consequence. With lower intensity, genetic gain becomes more sensitive to changes in d and the additional term in ΔF is also lower. Therefore, the balance will be more likely to favour $d = 1$ for a fixed N as ΔF decreases. Nevertheless, the change in the value of the optimum mating ratio from $d = 1$ to $d = 2$ in the circumstances considered here had a very small effect on genetic gain (Figure 4).

Computer simulation was used to check that optimal d is higher than 1 under some circumstances. One set of conditions for which predicted optimal d is 2 was considered ($N = 3000$ and $h^2_{(0)} = 0{\cdot}8$). For $N = 3000$, $h^2_{(0)} = 0{\cdot}8$ and $\lambda = 0$, the optimum N_s for fixed $d = 1$ is 20 and the optimum N_s for fixed $d = 2$ is 15. Predicted average gains over generations 5 to 14 are $P_C = 1{\cdot}3257$ for $N_s = 20$ and $d = 1$ and $P_C = 1{\cdot}3330$ for $N_s = 15$ and $d = 2$. Five thousand replicates were run for both schemes. The simulated average gains over generations 5 to 14 were 1·3296 for the scheme with $N_s = 20$ and $d = 1$ (standard error varied over generations from 0·0015 to 0·0016) and 1·3385 for the scheme with $N_s = 15$ and $d = 2$ (standard error varied over generation from 0·0014 to 0·0015). Thus, simulation results were remarkably close to those predicted showing that optimal d is 2 under some conditions as predicted by the full model used here. Although the difference in genetic gain when $d = 1$ or $d = 2$ was very small (again as predicted) it was significant.

The high reproductive capacity in females in fish species allows the number of dams to be less than the number of sires. The equation for predicting ΔF used here assumes $N_d > N_s$. An equivalent equation can be used for a situation where $N_s > N_d$, by replacing N_s by N_d, N_d by N_s and by defining d as N_s/N_d. This implies a dip in the response surface for mating ratio. In other words, for d defined as N_d/N_s, predicted ΔG is the same for $d = 1/2$ and $d = 2$ and higher than ΔG for $d = 1$. This may seen unusual but it has been shown here both through theory and through simulation.

Nicholas (1989) suggested that acceptable annual rates of inbreeding should be not higher than 0·5% and also proposed coefficients of variation of genetic gain ($CV(\Delta G)$) to be not higher than 5 to 10% after 10 years of selection. Under these conditions, genetic gains obtained by restricting the coefficient of variation of response to a satisfactory value are higher than that obtained by restricting inbreeding at an acceptable level, except for schemes of size $N <$ 1200 and $h^2_{(0)} = 0{\cdot}2$. The results presented are given per generation. Therefore, for species with generation interval of 1 year (e.g. tilapia), restricting the rate of inbreeding to acceptable levels automatically leads to acceptable coefficients of variation of response. These results agree with those obtained by Meuwissen and Woolliams (1994a; 1994b).

Although this paper deals with mass selection, the same procedure can be applied when selection is on an index including information on relatives or several traits. Equilibrium expressions (Bulmer effect) for genetic gain have been obtained for indices including some relatives (Gomez-Raya and Burnside, 1990) or several traits (Villanueva and Kennedy, 1990). Also, recurrent equations for predicting gain at each generation are available for different selection indices with different sources of information (Wray and Hill, 1989) and several traits (Villanueva *et al.*, 1993). These include indices which approximate best linear unbiased prediction (BLUP). Predictions of rates of inbreeding for populations under selection on an index including the individual record and full-sib and half-sib records have been already published (Wray *et al.*, 1994). Preliminary results for inbreeding under BLUP selection have also been provided (T. H. E. Meuwissen and J. A. Woolliams, unpublished

results). The situation is more complex in that, with information on relatives in the selection criterion, the weights given to different sources can be varied in order to maximize the objective function. However, many practical schemes are operated using mass selection and the optimization described here offers a proactive approach to the design of such breeding schemes.

Acknowledgements

This work was funded by the European Union and the Biotechnology and Biological Sciences Research Council (BBSRC). We thank Drs P. Amer and A. Caballero for many useful discussions and Drs P. Amer, G. Simm, B. Grundy and R. Thompson and the referee of the journal for useful comments on the manuscript.

References

Bentsen, H. B. and Gjerde, B. 1994. Design of fish breeding programs. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 19,* pp. 353-359.

Brisbane, J. R. and Gibson, J. P. 1994. Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 19,* pp. 135-138.

Bulmer, M. G. 1971. The effect of selection on genetic variability. *American Naturalist* **105:** 201-211.

Caballero, A., Santiago, E. and Toro, M. 1996a. Systems of mating to reduce inbreeding in selected populations. *Animal Science* **62:** 431-442.

Caballero, A., Wei, M. and Hill, W. G. 1996b. Survival rates of mutant genes under artificial selection using individual and family information. *Journal of Genetics* In press.

Gjerde, B., Gjøen, H. M. and Villanueva, B. 1996. Optimum designs for fish breeding programmes with constrained inbreeding. I. Mass selection for a normally distributed trait. *Livestock Production Science* In press.

Goddard, M. E. 1987. Policy of selecting bulls to breed bulls. *Animal Production* **44:** 29-38.

Gomez-Raya, L. and Burnside, E. B. 1990. The effect of repeated cycles of selection on genetic variance, heritability and response. *Theoretical and Applied Genetics* **79:** 568-574.

Grundy, B., Caballero, A., Santiago, E. and Hill, W. G. 1994. A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. *Animal Production* **59:** 465-468.

Hill, W. G. 1976. Order statistics of correlated variables and implications in genetic selection programmes. *Biometrics* **32:** 889-902.

Jódar, B. and López-Fanjul, C. 1977. Optimum proportions selected with unequal sex numbers. *Theoretical and Applied Genetics* **50:** 57-61.

Meuwissen, T. H. E. and Luo, Z. 1992. Computing inbreeding coefficients in large populations. *Genetics Selection Evolution* **24:** 305-313.

Meuwissen, T. H. E. and Woolliams, J. A. 1994a. Maximising genetic response in breeding schemes of dairy cattle with constraints on variance of response. *Journal of Dairy Science* **77:** 1905-1916.

Meuwissen, T. H. E. and Woolliams, J. A. 1994b. Response versus risk in breeding schemes. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 18,* pp. 236-243.

Meuwissen, T. H. E. and Woolliams, J. A. 1994c. Effective sizes of livestock populations to prevent a decline in fitness. *Theoretical and Applied Genetics* **89:** 1019-1026.

Nicholas, F. W. 1989. Incorporation of new reproductive technology in genetic improvement programmes. In *Evolution and animal breeding* (ed. W. G. Hill and F. C. MacKay), pp. 203-209. CAB International, Wallingford.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. 1992. *Numerical recipes in* FORTRAN: *the art of scentific computing. Second edition,* pp. 390-395. Cambridge University Press.

Santiago, E. and Caballero, A. 1995. Effective size of populations under selection. *Genetics* **139:** 1013-1030.

Stam, P. 1980. The distribution of the fraction of the genome identical by descendant in finite random mating populations. *Genetical Research* **35:** 131-155.

Toro, M. and Nieto, B. 1984. A simple method for increasing the response to artificial selection. *Genetical Research* **44:** 347-349.

Toro, M. and Perez-Enciso, M. 1990. Optimization of selection response under restricted inbreeding. *Genetics Selection Evolution* **22:** 93-107.

Verrier, E., Colleau, J. J. and Foulley, J. L. 1990. Predicting cumulated response to directional selection in finite panmictic populations. *Theoretical and Applied Genetics* **78:** 833-840.

Verrier, E., Colleau, J. J. and Foulley, J. L. 1993. Long-term effects of selection based on animal model BLUP in a finite population. *Theoretical and Applied Genetics* **87:** 446-454.

Villanueva, B. and Kennedy, B. W. 1990. Effect of selection on genetic parameters of correlated traits. *Theoretical and Applied Genetics* **80:** 746-752.

Villanueva, B., Woolliams, J. A. and Simm, G. 1994. Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics Selection Evolution* **26:** 517-535.

Villanueva, B., Wray, N. R. and Thompson. R. 1993. Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. *Animal Production* **57:** 1-13.

Woolliams, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49:** 1-14.

Woolliams, J. A. and Thompson, R. 1994. A theory of genetic contributions. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph, vol. 19*, pp. 127-134.

Woolliams, J. A., Wray, N. R. and Thompson, R. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research* **62:** 231-242.

Wray, N. R. and Goddard, M. E. 1994. Increasing long term response to selection. *Genetics Selection Evolution* **26:** 431-451.

Wray, N. R. and Hill, W. G. 1989. Asymptotic rates of response from index selection. *Animal Production* **49:** 217-227.

Wray, N. R., Woolliams, J. A. and Thompson, R. 1994. Prediction of rates of inbreeding in populations undergoing index selection. *Theoretical and Applied Genetics* **87:** 878-892.

(*Received* 11 *August* 1995—*Accepted* 5 *July* 1996)

Appendix

Approximation to genetic gain accounting for linkage disequilibrium and inbreeding

Following Woolliams and Thompson (1994) the rate of response to genetic selection in a generation can be seen as the expected crossproduct of the change in contribution of each ancestor with its Mendelian sampling terms, summed over generations. Consider $\Delta G_{(t)} = G_{(t)} - G_{(t-1)}$. Assuming that σ^2_P does not change, the Mendelian sampling variance for individuals born in generation $t-1$ is $(0{\cdot}5)\, h^2_{(0)}\, (1-\Delta F)^{t-3}$ since generation 3 is the first generation for which the sampling variance is reduced (generation 2 has the first inbred parents). Therefore the regression of the Mendelian sampling term (α) on the phenotype (P) is given by

$$\alpha = (0{\cdot}5)h^2_{(0)}\, (1-\Delta F)^{t-3}\sigma^{-2}_{P(L)}P + \varepsilon$$

where $\sigma^2_{P(L)}$ is the limiting value (Bulmer equilibrium) for the phenotypic variance. Thus, selection on P induces an increase in α among selected individuals by $(0{\cdot}5)h^2_{(0)}i(1-\Delta F)^{t-3}\sigma^{-1}_{P(L)}$ per unit of P.

The expected contributions of selected individuals of sex x to the next generation are equal assuming equal mating opportunities. However, in subsequent generations, the expected contributions will increase or decrease in a linear relationship with breeding value. The expected contribution of a selected ancestor $i(x)$ with breeding value $A_{i(x)}$ in generation t-j-1 changes by $(0{\cdot}25)N_x^{-1}\sigma^{-1}_{P(L)}\, ib^{j-1}A_{i(x)}$ (Woolliams *et al.*, 1993), where $b = (0{\cdot}5)(1 - kh^2_{(L)})$ and N_x is N_s for males and N_d for females. The expected crossproduct of the breeding value of an individual in generation t-j-1 with its Mendelian sampling term is $(0{\cdot}5)\sigma^{-1}_{P(L)}\, ib\, h^2_{(0)}\, (1-\Delta F)^{t-j-3}$. The value of h^2 in b should strictly be the appropriate for the generation, but it will be seen that little error but much simplification comes from this approximation. The total crossproducts of change in contributions and Mendelian sampling terms, summed over all ancestors in generation t-j-1 is therefore, $(0{\cdot}5)\sigma^{-1}_{P(L)}\, ib^j\, h^2_{(0)}\, (1-\Delta F)^{t-j-3}$. The terms decrease rapidly with t as $b < 0{\cdot}5$. Summing over generations,

$$\Delta G_{(t)} \approx \sum_{j=1}^{t-1} \frac{h^2_{(0)}i}{2\sigma_{P(L)}}\, b^j(1-\Delta F)^{t-j-3} + \frac{h^2_{(0)}i}{2\sigma_{P(L)}}\,(1-\Delta F)^{t-3} \quad \text{(A1)}.$$

Equation (A1) is modified slightly in that it assumes for generations 0 and 1 the Mendelian sampling terms are $(0{\cdot}5)h^2_{(0)}(1-\Delta F)^{-2}$ and $(0{\cdot}5)h^2_{(0)}(1-\Delta F)^{-1}$, respectively which are different from the true values of $(0{\cdot}5)h^2_{(0)}$. However, (i) the deviation is small and (ii) these are weighted by b^{t-1} and b^{t-2} which are also small values of t of 4 or more.

Equation (A1) is equivalent to

$$\Delta G_{(t)} \approx \sum_{j=0}^{t} \frac{h^2_{(0)}i}{2\sigma_{P(L)}}\, b^{t-j}(1-\Delta F)^{j-3}.$$

Here the fact that the base generation must account for the full variance and not only for its Mendelian term has been allowed for by adding a further generation of Mendelian terms. Again note the addition is weighted by b^t so is of diminishing importance. Thus,

$$\Delta G_{(t)} \approx \frac{h^2_{(0)}i}{2\sigma_{P(L)}}\, \frac{b^t}{(1-\Delta F)^3} \sum_{i=0}^{t} \frac{(1-\Delta F)^j}{b^j}$$

and defining $a = (1-\Delta F)$,

$$\Delta G_{(t)} \approx \frac{h^2_{(0)}i(a^{t+1} - b^{t+1})}{2\sigma_{P(L)}a^3(a-b)} \quad \text{(A2)}.$$

Equation (A2) has assumed $\sigma_{P(L)}$ remains constant. However, this approximation was found not to be sustainable for long. In generation 3,

$$\sigma^2_{P(3)} \approx \sigma^2_{P(L)} - (0{\cdot}5)\Delta F h^2_{(0)}$$

$$\approx \sigma^2_{P(L)}\left(1 - \frac{\Delta F h^2_{(0)}}{2\sigma^2_{P(L)}}\right)$$

and $$\sigma_{P(3)} \approx \sigma_{P(L)}\left(1 - \frac{\Delta F h^2_{(0)}}{4\sigma^2_{P(L)}}\right)$$

but after several further generations, the total genetic variance reduces by approximately $\Delta F\sigma^2_{A(L)}$ each generation. Therefore

$$\sigma^2_{P(t)} = \sigma^2_{P(t-1)} - \Delta F\sigma^2_{A(L)}$$

$$\approx \sigma^2_{P(t-1)}\,(1 - \Delta F h^2_{(L)}/\sigma^2_{P(L)})$$

so $$\sigma_{P(t)} \approx \sigma_{P(t-1)}\,[1 - (0{\cdot}5)\,\Delta F h^2_{(L)}\,\sigma^{-2}_{P(L)}].$$

This proportional reduction will be repeated in each subsequent generation and so, in generation $t-1$,

$$\sigma_{P(t)} \approx \sigma_{P(L)}\,[1 - (0{\cdot}5)\,\Delta F h^2_{(L)}\,\sigma^{-2}_{P(L)}]^{t-3}.$$

This term varies in the same manner as the Mendelian sampling variance and so $a = (1-\Delta F)$ may be replaced by a

$= (1 - \Delta F)[1 - (0{\cdot}5)\ \Delta F h^2_{(L)}\ \sigma^{-2}_{P(L)}]^{-1}$ in the above expressions. As for the Mendelian terms the errors induced in generation 1 and 2 are both small and weighted by terms of the order of b^t.

For t large, b^t is small since b is less than 0·5 and a is less than but close to 1. Therefore, for large t,

$$\Delta G_{(t)} \approx \frac{h^2_{(0)} i a^{t-2}}{2\sigma_{P(L)}\,(a-b)}$$

and for a very close to 1,

$$\Delta G_{(t)} \approx \frac{h^2_{(0)} i a^{t}}{2\sigma_{P(L)}\,(a-b)}$$

This approach using genetic contributions can be simply extended to include sib indices.

Paper 30

Anim. Prod. 1991, 53: 157-164

0003-3561/91/17950157$02-00

Utilization of the sex-determining region Y gene in beef cattle breeding schemes

S. C. Bishop and J. A. Woolliams
AFRC Institute of Animal Physiology and Genetics, Edinburgh Research Station, Roslin, Midlothian EH25 9PS

Abstract

In mammals 'maleness', i.e. the presence of testes, is thought to be controlled by a single gene on the Y chromosome. Recently, a candidate gene termed the SRY (sex-determining region Y) gene has been located. If the SRY gene is the gene causing maleness then a transgenic male with the SRY gene on an autosome would produce a greater proportion of male offspring than a normal male. This would be advantageous in situations where male offspring are more valuable than females. Such transgenic males have a reduced probability of propagating their genotype and an effort has to be made to avoid their extinction. This is at the cost of genetic progress which must be made to enable the transgenics to remain competitive with normal males.

In a simulated beef cattle breeding scheme if half of the annual matings were made to transgenics then after 15 years of selection the transgenic males fell the equivalent of 2·6 years of selection behind males in a traditional herd. If all matings were made to transgenics they fell over 9 years behind. Selection for lean food conversion ratio was considered as an example. After 15 years of selection the gain in biological efficiency from more male offspring outweighed the loss from reduced genetic progress only when more than 0·5 of the bulls used in the breeding scheme were normal males. In practice, the difficulty of maintaining a small population of transgenic males along with other costs not included in the calculations suggest that breeding schemes in beef cattle with an SRY transgene would not be practicable without further technology.

Keywords: *beef cattle, sex ratio, selection, SRY, transgenic animals.*

Introduction

A gene, termed SRY (sex-determining region Y), located on the Y chromosome, has recently been isolated and described by Sinclair, Berta, Palmer, Ross Hawkins, Griffiths, Smith, Foster, Frischauf, Lovell-Badge and Goodfellow (1990). This gene is currently considered to be a candidate for the testes-determining gene found on the Y chromosome of normal males. In normal mammalian sexual development ovarian development is considered to be the 'default pathway' and the testes-determining gene is thought to be the factor which determines 'maleness', or presence of testes, in individuals possessing the Y chromosome. Evidence supporting this claim includes the fact that males have been observed with an XX genotype but with a fragment of the Y chromosome containing the SRY gene attached to one of the X chromosomes (Sinclair *et al.*, 1990), and female mice with an XY genotype have been found with the same region missing from their Y chromosome (Gubbay, Collingnon, Koopman, Capel, Economou, Münsterberg, Vivian, Goodfellow and Lovell-Badge, 1990).

Identification and exploitation of this 'male gene' is thought to have potential benefits in agriculture, especially in animal production systems where the male offspring is more valuable than the female. Perhaps the simplest application of the SRY gene would be to create a transgenic male carrying the SRY gene on an autosome. This animal would have an altered sex ratio amongst its offspring, producing more male offspring than a normal male. This type of transgenic animal differs from forms normally considered in the agricultural context insofar as the gene transferred is not one directly controlling a production trait, e.g. growth hormone, but rather it indirectly alters the expression of production traits by changing the reproductive characteristics of the animal.

An obvious application of the SRY gene is in beef-cattle production systems using terminal-sire breeds where offspring of both sexes are marketed for the same trait, but where the male calf attracts a financial premium. This financial premium will confer an advantage to transgenic sires, however this benefit will be cancelled out by a normal sire with a sufficiently large superiority in its breeding value for the marketable trait. Utilization of the transgenic males therefore has to be considered within a selection framework because unless the transgenic males make genetic progress they will be superceded quickly by normal males which are genetically superior for this trait. This paper explores, by means of simulation, the possible utilization and benefits from transgenic animals containing the SRY gene in the context of beef-cattle breeding.

Material and methods

It is assumed that male transgenic animals are available with the SRY gene on an autosome, and that this gene is functional and is necessary and sufficient for maleness. It is also assumed that these animals can be easily distinguished from normal males before counting their offspring. This can be achieved by DNA-typing if the SRY transgene is tagged before insertion. The genotype of this animal will be designated as [XY, A_oA_m], where A_o is the normal autosome, A_m is the genetically altered autosome, and X and Y are the normal sex chromosomes.

Mating this animal to a normal female [XX, A_oA_o] results in equal proportions of the following genotypes: [XX, A_oA_o], [XX, A_mA_o] [XY, A_oA_o] and [XY, A_mA_o]. These four genotypes will be defined as female, pseudo male, normal male and super male, respectively, giving a phenotypic ratio of one female to three males.

Of the three male genotypes only the super males [XY, A_mA_o] will sire further super males because normal males [XY, A_oA_o] can sire only females and normal males, and pseudo males [XX, A_mA_o] can sire only females and pseudo males. Therefore, in a genetic sense super males are less fit than normal and pseudo males because they have only a 0·25 chance of propagating their own genotype whereas normal and pseudo males have a 0·50 chance of propagating their genotypes. As a result, a breeding scheme where selection is only on genetic merit for the trait of interest, regardless of whether the animals are transgenic or not, will lead to extinction of super males because a smaller number will be available for selection each successive generation. A positive effort must therefore be made to maintain super males in the herd. The implications of maintaining different proportions of super males within a nucleus herd are investigated in the simulations described below.

Structure of breeding scheme

A nucleus breeding scheme for beef cattle was considered with 1000 cows expected to calve each year, and one calf per calving. The objective of the nucleus was to produce bulls for use as terminal sires. It was assumed that only super males and normal males were used for breeding within the nucleus and that pseudo males were discarded, however the validity and impact of this assumption were considered. A total of M bulls were mated to the cows, with each bull being allocated the same number of cows whether it was a super male or normal male. The proportion of normal bulls used was Q. Under these assumptions the proportions of genotypes amongst the calves born will be 0·25 (1 + Q) for females and normal males and 0·25 (1 − Q) for super and pseudo males. The proportion of males selected for breeding, p, was defined as the proportion of normal males that would be selected if $Q = 1$, thus in this example $p = M/500$. All normal and super male calves were assumed to be eligible for selection (i.e. no mortality) and selection was on the individual's own phenotype. Offspring of selected bulls were born at 2 years of age. Since the selection intensity of females was small, particularly for $Q = 0$, it was assumed that for economic reasons females were selected at birth on the average of their parental phenotypes. All females were assumed to survive to 14 months of age by which time their individual phenotypes were measured. Although this was not used for selection of females it was available for use in assortative mating. After 14 months of age a random fractional loss of 0·15 occurred in females by first calving at 2 years of age, and again, annually, between each successive calving. Given the rates of loss described and the constraint of 1000 expected calvings, the number of females selected at birth and the maximum number of calvings per cow was determined for each value of Q and M to maximize the average genetic merit of all offspring born (prior to selection) in the first generation. The number of female calves selected and the maximum number of calvings per cow are shown in Table 1 for various values of Q and M. For low Q values it can be seen that it is difficult to generate sufficient female calves to maintain the nucleus at 1000 breeding cows.

Selection strategies

Various alternative strategies were considered.

Strategy I. Cows were randomly allocated to bulls irrespective of whether they were normal or super males (IA); variations were considered in which

Table 1 *The number of female calves selected and the maximum number of calvings per cow for various numbers of sires used and proportions of transgenics*

	No. of sires used (*M*)					
	25			50		
Q†	0	0·5	1·0	0	0·5	1·0
No. of female calves selected	241	335	435	235	329	388
(No. of females born)	(250)	(375)	(500)	(250)	(375)	(500)
Maximum number of calvings‡ per cow	9	5	4	9	5	4

† Proportion of sires which are normal males, pseudo males not considered.
‡ Maximum number of calvings is the number necessary for some cows in order to maintain 1000 calvings per year.

pseudo males were treated as normal males (IB), or all calves born to super males except super males being excluded from breeding (IC).

Strategy II. The best cows were preferentially mated to super males (IIA) or to normal males (IIB), with cows being randomly allocated to bulls within the normal and super male mating groups. For $Q = 0$ or $Q = 1$ strategies I and II are equivalent.

Strategy III. A more realistic situation where only one super male, i.e. a newly created transgenic, is available in each of the first 2 years and the single bull was mated to 1000 $(1 - Q)$ cows so that the number of calves of each genotype remained constant over generations. The ancestral super males are assumed to be of either average genetic merit (IIIA) or 1 phenotypic s.d. above the initial mean (IIIB).

In strategies I and II the initial population was assumed to be drawn at random in the appropriate proportions from infinite populations of females, normal males and super males with the same mean, phenotypic variance (σ_p^2) and additive genetic variance (σ_a^2). Strategy IA was simulated for $p = 0{\cdot}05$, 0·1 and 0·2, for $h^2 = 0{\cdot}25$ and 0·5 and for $Q = 0$, 0·1, 0·2,.....0·9, 0·95, 0·98 and 1·0. All other strategies were simulated over the same range of values for Q but for $p = 0{\cdot}1$ and $h^2 = 0{\cdot}5$ only.

Simulation technique
Initially all animals were assigned a breeding value randomly drawn from a normal distribution of mean 0 and variance h^2 ($N(0, h^2)$). The breeding value of offspring were derived by the addition of a Mendelian sampling term randomly drawn from $N(0, \frac{1}{2}h^2)$ to the average of the parental genotypes. The phenotype of all animals was derived by the addition of an environmental deviation randomly drawn from $N(0, 1 - h^2)$. Selection was undertaken for 15 years and each strategy and each combination of values for h^2, p and Q was replicated 50 times.

Results and discussion

Analysis of simulations
The results of different values of p and h^2 were qualitatively very similar, so only the results for $p = 0{\cdot}1$ and $h^2 = 0{\cdot}5$ are presented. Important comparisons are those between normal and super male herdmates, and between super males in the model herd and normal males in a traditional herd ($Q = 1$ and random mating).

Figure 1 shows the average genetic merit of normal and super male calves for strategy IA after 15 years of selection. Female calves have the same average genetic merit as normal male calves because they are produced by normal and super males in the same proportions as normal males. Likewise, pseudo male calves have the same expected genetic merit as super male calves. Super males were genetically inferior to their normal male herdmates for all Q and the divergence increased with Q.

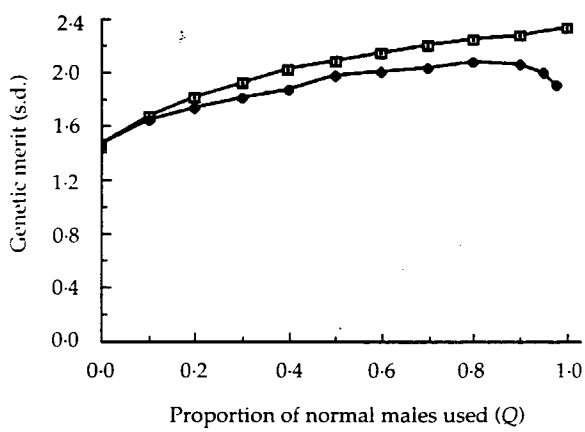


Figure 1 Genetic merit of normal and super males after 15 years of selection. $Q = 1$ corresponds to a breeding scheme in a traditional herd; □ normal males; ◆ super males.

Table 2 *The genetic lag after 15 years of selection between the mean genetic merit of normal males when no super males are used and normal and super males in the model herd, for different values of Q*

Q:	0·00	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	0·95	0·98
Normal males:												
s.d. s	0·89	0·68	0·53	0·43	0·33	0·27	0·19	0·13	0·09	0·06	0·05	0·00
years†	9·10	5·87	4·23	3·31	2·37	1·83	1·26	0·84	0·56	0·35	0·31	0·00
Super males:												
s.d. s	0·89	0·72	0·61	0·54	0·45	0·38	0·32	0·30	0·25	0·26	0·29	0·37
years	9·10	6·30	4·91	4·13	3·21	2·63	2·15	1·94	1·56	1·63	1·73	2·22

† Years is the number of years of genetic progress in the model herd.

The genetic inferiority of super males would be expected from an analysis of the initial selection proportions (see **Appendix** 1). In the first generation the selection proportion for normal males is $2pQ/(1 + Q)$, i.e. a strictly increasing function of Q, which achieves its maximum value of p when $Q = 1$, whereas for super males it is $2p$ for all values of Q. This initial difference in selection proportions is a reflexion of the relative fitness of the normal and super male genotypes and it generates the initial genetic superiority of females and normal males over super males. From the first generation onwards the different genotypes have different means and the different types of mating makes deterministic analysis of subsequent generations difficult, however, the pattern of annual genetic gains for normal and super males derived from simulation is shown in Figure 2. The initial genetic difference (lag) between super and normal males widened and then stabilized after about 7 years with both groups making genetic progress at the same rate. In the example shown in Figure 2 for $Q = 0·5$ this lag was approximately 0·8 years of genetic progress. An estimate of this lag under scheme IC can be obtained easily where the selection proportion for normal males remains at p. The difference in selection intensity between normal and super males for this scheme is $\delta = [i(500pQ, 500Q) - i(500,p(1 - Q), 250(1-Q)]$, where $i(n_1, n_2)$ is the standardized selection intensity when selecting n_1 animals from a total of n_2. The lag will increase until the genetic difference is $h^2\delta(\sigma_p$ units).

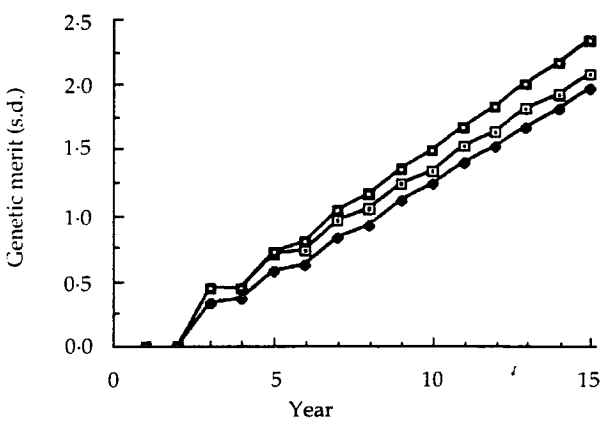


Figure 2 Response to selection for $Q = 0·5$ compared with response in a traditional herd ($Q = 1$); ▣ traditional herd; ◫ normal males, $Q = 0·5$; ◈ super males, $Q = 0·5$.

Also shown in Figure 2 is the comparison of animals selected in a traditional herd with those in a herd with $Q = 0·5$. Initially normal males in the two herds are comparable, however the super males are always inferior and this gap widens with time. A summary of the genetic lags after 15 years of selection between the normal and super male herdmates and a traditional herd, for different values of Q is shown in Table 2. The lag is expressed in standard deviation units and also in terms of years of selection, thus indicating how many years it would take the normal and super male to reach the genetic level of a herd selected without any super males. For small Q, there was a very large lag but this decreased as Q increased. The minimum genetic distance between super males and a traditional herd was 1·56 years, when $Q = 0·8$. The lag between males in the model herd and a traditional herd would increase with time since more rapid annual rates of progress are made when $Q = 1$.

Figure 3 shows the between replicate variation in responses to selection after 15 years for the normal and super males, calculated as the empirical standard deviations of the 50 replicate means. These may also give an indication of the inbreeding rate in the different groups of the selection herd; the variance of genetic change in a trait is directly related to inbreeding when selection is random with respect to that trait (Hill, 1977). The standard deviations for the normal male means were almost constant for all values of Q, giving coefficients of variation decreasing from 0·05 at $Q = 0$ to 0·03 at $Q = 1$. The standard deviations for the super male means increased rapidly beyond $Q = 0·3$ because the number of sires giving rise to super male calves was directly proportional to $(1 - Q)$ and hence became

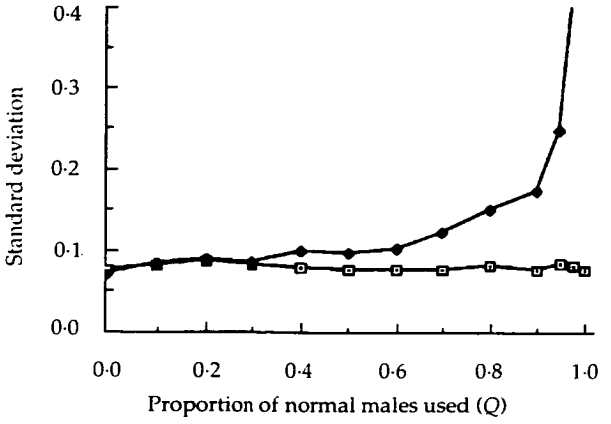


Figure 3 Standard deviation of replicate means for normal and super males after 15 years of selection. $Q = 1$ corresponds to a breeding scheme in a traditional herd; ▫ normal males; ◆ super males.

very small for large Q. The coefficient of variation for super males rose to 0·11 at $Q = 0{\cdot}95$ and 0·21 at $Q = 0{\cdot}98$.

Using strategy IIA, i.e. preferentially mating the best cows to super males, resulted in super males having greater genetic merit than their normal male herdmates (Figure 4a) but this was at a cost to the rate of genetic progress of normal males and females and the model herd as a whole. This is because the normal males, which have the greatest selection differential (and hence potential for making genetic progress) and who provide most of the female replacements are mated to the worst females. After 15 years of selection using strategy IIA super males are genetically superior to normal males in a traditional breeding scheme, for Q values greater than 0·6. They soon fall behind the traditional scheme, however, e.g. at year 21 for $Q = 0{\cdot}7$, and the lag increases with time.

Strategy IIB, i.e. preferentially mating the best cows to normal males, had the opposite effect to IIA (Figure 4b). For $Q > 0{\cdot}4$ the rate of genetic progress of the normal males and females was greater than in a traditional herd, but by year 15 the super males had fallen far behind normal males from a traditional herd. For $Q = 0{\cdot}7$ this genetic lag was 9 years after 15 years of selection. Because of the faster rate of progress of females and normal males, however, this gap gradually closes and in the very long-term super males will actually overtake normal males in a traditional herd. For $Q = 0{\cdot}7$ simulation showed this to occur at year 77.

Figure 5 shows results from strategy III, the 'realistic' situation of starting with only one transgenic male

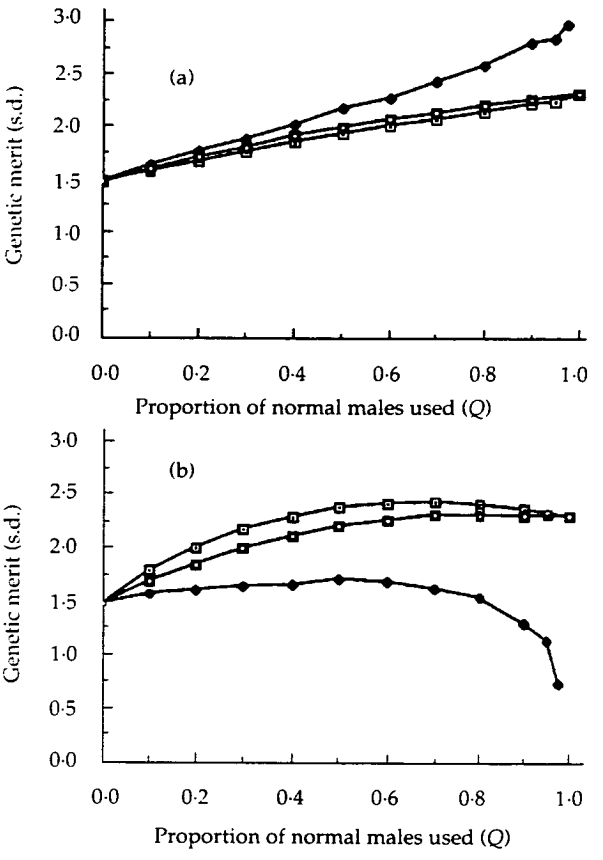


Figure 4 Effect of mating the best cows to (a) super males or (b) normal males on genetic merit after 15 years. $Q = 1$ corresponds to a breeding scheme in a traditional herd; ▫ normal males; ◆ super males; ■ weighted mean.

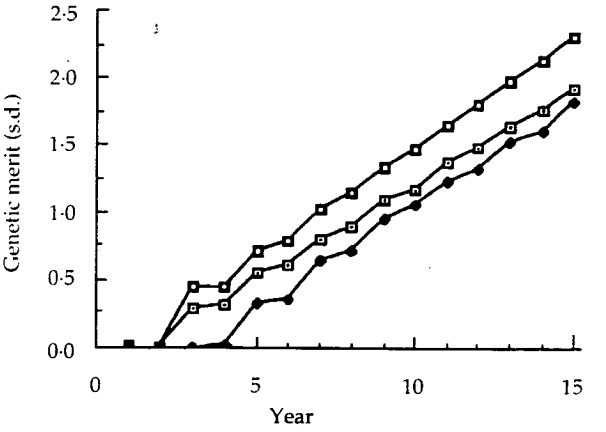


Figure 5 Response to selection for $Q = 0{\cdot}5$, with one initial super male of average genetic merit, compared with response in a traditional herd; ■ traditional herd; ▫ normal males, $Q = 0{\cdot}5$; ◆ super males, $Q = 0{\cdot}5$.

Table 3 *The genetic merit after 15 years of selection for strategies IA, IB and IC† (results are for Q = 0·5)*

Genetic merit	IA	IB	IC
Normal males	2·08	2·09	1·83
s.d.	0·078	0·078	0·099
Super males	1·97	2·00	1·68
s.d.	0·098	0·105	0·107
s.e.d.	0·012	0·011	0·016

† See **Material and methods.**

animal in each of the first 2 years, with these animals being of average genetic merit. Increasing the initial genetic merit of the super males increases the genetic merit of super males born in the first few years, and it also increases the genetic merit of the normal males, but subsequently the rate of genetic progress and genetic lag between the two genotypes stabilize to those for strategy IA. When the initial super males had a genetic superiority of one phenotypic standard deviation the genetic merit of the super male calves was equal to the normal male calves for the first 2 years, but by year 5 the responses to selection were almost identical to those shown in Figure 2. This was equivalent to a genetic lift of approximately 1 year over the animals shown in Figure 5. The initial genotype of the transgenic animals will therefore have a large effect on decisions made in the first few years but not in the longer term.

A comparison of strategies IA, IB (including pseudo males) and IC (rejecting all offspring born to super males except super males) is shown in Table 3. Adopting strategy IB made a very small increase in genetic gain compared with IA, however, in practice this would be at a cost of extra testing facilities. For small Q values, it was impossible to maintain the herd size, given the survival rates assumed, using strategy IC.

An alternative strategy for operating a breeding scheme using transgenic males would be to assign the super males a phenotypic bonus in relation to the advantage they have over normal males as terminal sires. A large bonus would ensure that only super males are selected (equivalent to $Q = 0$), whereas a bonus of zero would quickly result in their extinction. With intermediate levels of bonus an equilibrium Q level is reached, however this equilibrium is very unstable and effort has to be made to ensure that super males remain in the herd. This strategy therefore reverts to those already considered where a specific Q is maintained.

General considerations

The crucial assumption made in this paper is that additional sex chromosomes can be created which are inherited independently of the X and Y chromosomes and that, unlike abnormalities of the sex chromosomes documented in domestic animals (see Halnan, 1975), these chromosomes have no deleterious effects on fertility or performance. Instead of creating extra sex chromosomes the X chromosome could perhaps be targeted for insertion of the SRY gene. This would create an animal with effectively a YY genotype. This animal would produce all male offspring, however these offspring would all have the reproductive properties of normal males. After one generation of mating, therefore, the YY genotype would be lost to the breeding scheme.

The results must be put in context of the economic value of the trait or index under selection. The results given for each value of Q assume that the cow age structure has been optimized for female selection, given the number of females available for selection. This will be difficult to achieve in practice and hence the examples will tend to overestimate rates of genetic gain possible when super males are used. Moving from the optimal situations, e.g. not altering the cow age structure when Q is altered, would decrease the overall rate of genetic gain and hence put the herd at more of a disadvantage. The values for genetic lags in Table 2 should therefore be considered minimum values. However, suboptimality has little effect on the comparison of normal and super males within the model herd.

A suitable trait for selection is lean food conversion ratio (LFCR), estimated as food intake/(weight gain × predicted killing-out proportion × predicted carcass lean proportion). This trait is simply a prediction of the efficiency of growth of saleable carcass lean, and in a selection experiment with Hereford cattle described by Mrode, Thompson and Smith (1990) it responded to selection with a realized heritability of 0·4, having a mean of 17·4 (kg/kg) and a phenotypic standard deviation (σ_p) of 2·47 (kg/kg). The question is: would a breeding scheme incorporating super males and selecting for LFCR be competitive with a traditional scheme selecting for LFCR?

The most favourable result for super males using strategy IA was when $Q = 0·8$, where super males are only 1·56 years or $0·25\,\sigma_p$ behind males from a traditional herd after 15 years of selection. The genetic difference in the offspring will be 0·78 years or $0·125\,\sigma_p$. A super male produces one extra male calf in four, so for it to be advantageous to a farmer the difference in value between male and female calves must be four times this value, i.e. $0·50\,\sigma_p$ or 1·23 (kg/kg). This is equivalent to a minimum proportional benefit in favour of males of 0·07. The performance of Hereford × Friesian progeny of the

animals described by Mrode *et al.* (1990) has been evaluated in both steers and heifers, and the proportional advantage to steers in LFCR was found to be 0·077 (Bishop, Broadbent, Kay, Rigby and Fisher, 1992). In agreement with this value, Taylor (1982) calculated the proportional benefit of steers over heifers in overall efficiency to be 0·07, and the benefit of bull calves over heifers to be 0·13. None of these values allows for extra costs that would be incurred by super males through a greater incidence of dystocia or a premium charged for the semen, and therefore they must be seen as favourable towards super males. Given these reservations, it may be beneficial to use super male from a breeding scheme operating with $Q = 0{\cdot}08$.

A nucleus scheme would almost certainly not operate at $Q = 0{\cdot}8$, however, as this results in super male to normal male calves in the ratio of only 1 : 9. The risks are producing insufficient super males to cover the costs of creating the transgenics in the first instance and the costs of subsequent identification of super males, and it may also create problems in the maintenance of a population with a sire line of small and restricted size. A compromise would have to be made whereby a greater proportion of super males was used in the nucleus, to generate more super male bulls for commercial use, at a cost in terms of genetic gain. The calculations of the benefits of super males are also dependent on the time horizon chosen because a breeding scheme incorporating super males continually falls further behind a traditional scheme. The longer the time horizon chosen the less the possible benefits from super males.

This paper has investigated the application of the SRY gene at the simplest level. Increasing the degree of technology used in a breeding scheme may increase the possible benefits, but at an increased cost. For example, incorporation of SRY gene into a multiple ovulation and embryo transfer (MOET) scheme with embryo sexing raises further possibilities and may be of benefit. In the model scheme described pseudo males have no genetic benefit. Using MOET, and sexing prior to transfer would allow pseudo male embryos to be removed and replaced by whichever sex is least represented or genetically the most beneficial. Such a scheme might, in theory, be able to become competitive by using only super males and females. Of considerable further complexity and requiring a greater degree of commercial control is the construction of conditional promoters which would regulate the switching on of the gene by the farmer. The value of the transgene is perhaps most readily appreciated in beef cattle breeding. In dairy cattle breeding the female calf is of more value to the farmer and the simple transgenic male would not be appropriate. In this context a pseudo male with the SRY gene activated by a conditional promoter may be beneficial as it could, theoretically, sire only female calves.

In conclusion, with development of transgenic technology it may become possible to breed cattle with a copy of the SRY gene on an autosome, thus changing the sex ratio of the offspring from 1 : 1 to 3 : 1 in favour of males. However, inclusion of the SRY gene in a breeding scheme will be at the cost of genetic gain and this along with the additional costs of creating and identifying the transgenic animals, makes it unlikely that without the continued support of other embryo techniques such a breeding scheme could compete with a conventional breeding scheme.

References

Bishop, S. C., Broadbent, J. S., Kay, R. M., Rigby, I. and Fisher, A. V. 1992. The performance of Hereford × Friesian offspring of bulls selected for lean growth rate and lean food conversion efficiency. *Animal Production*. In press.

Gubbay, J., Collingnon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian N., Goodfellow, P. N. and Lovell-Badge, R. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature, London* **346:** 245-250.

Halnan, C. R. E. 1975. Chromosomes of cattle: present clinical status and promise. *Veterinary Record* **96:** 148-151.

Hill, W. G. 1977. Variation in response to selection. *Proceedings of the international conference on quantitative genetics* (eds. E. Pollak, O. Kempthorne, and I. B. Bailey), pp. 343-365.

Mrode, R. A., Thompson, R. and Smith, C. 1990. Selection for lean growth rate and the efficiency of lean growth in Hereford cattle. 1. Direct responses to selection. *Animal Production* **51:** 23-34.

Sinclair, A. H., Berta, P., Palmer, M. S., Ross Hawkins, J., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A-M, Lovell-Badge, R. and Goodfellow, P. N. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature, London* **346:** 240-244.

Taylor, St C. S. 1982. Theory of growth and feed efficiency in relation to maturity in body weight. *Proceedings of the 2nd world congress in genetics applied to livestock production*, Vol. V, pp. 218-230.

(*Received* 10 *January* 1991—*Accepted* 6 *April* 1991)

Appendix 1

Assume for $Q = 1$, M represents a proportion p of the normal males born. When Q is varied, the number of normal males required for breeding changes by a factor of Q but the number of normal males born changes by a factor $\frac{1}{2}(1 + Q)$. Thus the proportion of normal males selected is $2pQ/(1 + Q)$, and consequently as Q increases so selection proportion increases and intensity decreases.

In contrast, when $Q = 0$, to maintain the same ratio a proportion $2p$ of super males need to be selected (since only $\frac{1}{4}$ of their offspring are super males) and when Q is varied both the number required and the number born change by a factor $(1 - Q)$. Therefore, the proportion of super males selected remains constant at $2p$. Hence the selection proportion and intensity for super males are independent of Q.

Paper 31

Animal Science 1995, **61**: 177-187

1357-7298/95/33560177$20·00

Genetic contributions of Finnish Ayrshire bulls over four generations

J. A. Woolliams† and E. A. Mäntysaari

Agricultural Research Centre, Department of Animal Breeding, SF-31600, Jokioinen, Finland

Abstract

The long-term genetic contributions were calculated for 219 Finnish Ayrshire bulls born between 1958 and 1964 to 707 Finnish Ayrshire bulls made available for artificial insemination and born between 1986 and 1988. Three strategies were employed: (i) using all known pedigree information; (ii) ignoring information on the dam of females; (iii) only using information on sires. Expected contributions were calculated using gene flow matrices.

The contributions from strategies 1, 2 and 3 were only 0·6 (1 and 2) or 0·7 (strategy 3) of those expected. The causes of this shortfall for strategies 2 and 3 were identified as (i) the use of an imported sire and (ii) generation skipping. For strategy 1, 0·2 of the expected pathways remained unaccounted for and were ascribed to missing pedigree information.

Of the 219 ancestors, only 86 made positive contributions to the descendants. Only 10 ancestors made contributions more than the average, and one bull accounted for 0·3 of all pathways traced on strategy 2. There was general agreement in the relative contributions of individual bulls when assessed using the three strategies.

The rate of inbreeding (ΔF) estimated by regression from 1974 to 1988 and using known pedigrees was 0·0018 per year and the average coefficients of additive genetic relationship among cohorts was increasing by 0·0030 per year. ΔF was estimated using the contributions calculated by strategies 1, 2 and 3 to be 0·0147, 0·0151 and 0·0125 per generation respectively. These were converted into rates per year by assuming a generation interval of 6·5 years taken from both published and new information on generation intervals in the Finnish Ayrshire population. This gave annual rates of 0·0023, 0·0023 and 0·0019. The estimates from strategy 3 were obtained without the use of any pedigree information pertaining to dams.

Keywords: *cattle, Finnish Ayrshire bulls, gene flow, inbreeding.*

Introduction

Breeding schemes are not only concerned with the improvement of genetic merit but also with the protection of the genetic base during this process. The latter has customarily been measured by the rate of inbreeding (ΔF), since (i) it measures the rate of loss of variation and, (ii) it measures the rate of loss of heterogeneity at particular loci thereby giving an indication of the potential build up of unrecognized, possibly lethal recessives. However, it is often difficult to measure the rate of inbreeding; its calculation requires pedigrees going back to a base generation and incomplete pedigrees result in underestimation. A second problem is that few cattle populations remain closed for the considerable periods of time that are required for accurate estimation of the rate of inbreeding. The effect of an importation is to reduce immediately inbreeding and makes the apparent rate difficult to interpret. If importation is not an integral part of the breeding scheme this effect then makes the use of ΔF inadequate as a measure of risk in the practices of the breeding scheme before, during and after importation.

Wray and Thompson (1990) showed that an alternative approach to estimate inbreeding was to consider the long-term contributions of ancestors to

† On leave from Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS.

Table 1 *Mean generation intervals along the four genetic pathways for Finnish Ayrshire cattle between the years 1961 to 1966 (Lindström, 1969), 1970 to 1977 (Lindström, 1978) and 1989 (Finnish Animal Breeding Association)*

	Generation interval (years)						
	1961-1966			1970-1977			1989
Pathway	Mean	s.e.	Range	Mean	s.e.	Range	Mean
Bulls to breed bulls‡	9·29	0·11	8·86-9·52	8·96	0·08	8·01-9·69	8·00
Bulls to breed cows§	6·23	0·06	6·01-6·89	6·40	0·10	5·75-7·13	5·40
Cows to breed bulls	8·67	0·12	8·20-8·87	7·41	0·05	6·65-9·23	6·30
Cows to breed cows	4·87	0·05		4·50			4·20
Mean	7·27			6·82			5·98

† The range is over the mean values for each year of the period.
‡ Estimated from the parentage of young bulls entering AI service in this period.
§ Estimated by use during the period.

their descendants. This method has an advantage in that the authors were able to show that the relationships of male ancestors and female ancestors with descendants could be considered separately. In practical cattle breeding, information on bull ancestors is invariably more complete than on cow ancestors. Furthermore, whereas identities by descent increase with generations, long-term contributions are largely determined by the breeding pattern of the initial generations. A complication not cosidered by either Wray and Thompson (1990) or Woolliams, Wray and Thompson (1993) is the effect of overlapping generations on the approach.

Therefore it is appropriate to examine the approach of estimation of the long-term contributions of ancestors in an existing cattle population as a means of describing the risk attached to the breeding scheme. The Finnish Ayrshire population has had few but notable importations, has made good selection progress (0·9% per year, Mäntysaari and Stranden, 1991), and has been relatively well recorded, with the proportion of recorded herds now approaching 0·6. Furthermore, Lindström and Maijala (1971) conducted an extensive study of inbreeding and ancestral contributions on this population with data from bulls up to 1963 and cows up to 1965, and therefore it now seems timely to update this study. Finally, but by no means unimportantly, the identity of bulls making a large impact on the Ayrshire population is of general interest to all cattle breeders as well as being of specific importance to Ayrshire breeders for strategic planning. Therefore contributions were studied from all individuals in the cohorts of bulls born between 1958 and 1964 to the present population by bulls born between 1986 and 1988 and made available for artificial insemination (AI).

Methods

This section will include information about the data, programming methods and techniques of analysis, however, since many of the parameters used are designed to apply to the Finnish Ayrshire breed, a brief description and review of published information on this breed will also be given.

Finnish Ayrshire population and breeding

The Finnish Ayrshire population was originally developed from 1562 Ayrshire animals from Scotland and 303 from Sweden (Simonen, 1950). The period of interest for this study starts when the breed was already well established and considers bulls born in 1958 up to those born in 1988. During this time the total Finnish cow population has decreased from 1 134 889 cows to 534 600 cows, and the number of recorded cows from 314 649 to 293 576. During the same time the Ayrshire breed has gained a dominant position: in 1958 its share of milk recorded cows was 47·6% (150 075 cows) and in 1988 80·0% (235 162 cows). Lindström (1969 and 1978) provided considerable information as to the breeding practices operating in the period around 1958. Table 1 shows the generation intervals from 1960 to 1966 (Lindström, 1969) and from 1970 to 1977 (Lindström, 1978). The generation interval decreased from 7·27 years to 6·89 years over the combined period of the studies. Since that period an improved testing system has led to a further decrease in generation interval, which is currently close to 6 years (J. Juga, Finnish Animal Breeding Association; Table 1). During the period studied semen from Canadian and Norwegian bulls was imported. The year of birth of these bulls varied from 1965 to 1968 and also between 1986 and 1988. The Finnish Animal Breeding Association is responsible for the recording of pedigree information and for providing regular

genetic evaluations of the herd, but milk recording is the responsibility of the Association of Rural Advisory Centres.

Genetic methods

The pedigree database for the Finnish Ayrshire that was used contained a total of 978 548 records dating back to 1950. As a preliminary examination of the data the inbreeding for each cohort of bulls born between 1950 and 1988 was estimated together with the mean additive genetic relationship among all pairs of bulls in the same cohort. This was achieved using a version of the 'recursive subroutine getcoeff' (Miglior, Szkotnicki and Burnside, 1992). It was programmed in XL Fortran and an apparent error in the published version was corrected; where diagonal elements of the relationship matrix were being obtained i.e. the relationship of an animal *i* with itself, the published version calculated [1 + (getcoeff(*i*,isire) + getcoeff (*i*,idam))/2] and this was changed to (1 + getcoeff(isire,idam)/2), where *i*sire and *i*dam are the sire and dam of animal *i*.

An additional parameter was added to the subroutine to count the number of unknown ancestors encountered in the pedigree search. If the relationship between the oldest animal in the database and individual *i* was calculated, then the strategy employed by Miglior *et al.* (1992) will lead to the counting of the total number of distinct paths leading to an unknown ancestor in the pedigree. If an animal *j* with unknown parents occurred in *n* branches of *i*s pedigree then each of the unknown parents of *j* would be counted *n* times. As an example, consider discrete generations with pedigree information known apart from the parents of generation 1. Thus, with this definition, the number of unknown parents for an individual in generation 1 is 2, and for an individual in generation *t* is 2^t, irrespective of pedigree. Thus if N is the number of unknown parents, $\log_2 N$ is the number of generations of known pedigree. This definition was applied to the Ayrshire pedigrees: for each bull N was calculated and converted to 'discrete generation equivalents' (T) of pedigree information by $T = \log_2 N$. The reciprocal of the slope of the regression of T on year is an indication of the time taken for the number of ancestors to double but is not used in this paper as a measure of generation interval. In genetics the generation is defined as the average age of the parents when their replacements are born, and for overlapping generations with unequal intervals for males and females, these two definitions are *not* equivalent.

The primary objective was to obtain the long-term contributions of bulls of a single early generation on the most recent cohorts of bulls. The long-term contribution of an ancestor was defined by Wray and Thompson (1990) as the average genetic relationship of an ancestor with descendants many generations later (ignoring any relationship through pathways involving the parents of the ancestor) multiplied by the number of breeding individuals of both sexes per generation. Ultimately, over many generations, in a thoroughly mixed population an ancestor has the same relationship with all descendants but this will differ between ancestors. In this study the total number of breeding individuals is difficult to define and so the 'contribution' will be defined without this scaling factor and so is simply the proportion of all pathways leading back from descendants to ancestors that pass through the individual ancestor (an alternative terminology using 'average relationship' is potentially confusing since what is intended is not precisely the additive genetic coefficient of relationship). With discrete generations and with complete information the sum of the contributions over all ancestors in a single generation of each sex should be 0·5.

From the previous work on generation intervals, the early generation was chosen to be bulls born between the years 1958 to 1964 inclusive and these will henceforth be termed 'ancestors'. The most recent complete cohorts were bulls born in the years 1986 to 1988 and these will be termed 'descendants'. The information required was the contribution of each ancestor with the descendants which involves ignoring all relationships through pathways that pre-date the ancestor. This needed a further modification of the getcoeff subroutine in which individuals that become involved in the recursive procedure but which pre-date the ancestor are treated as unknown. In this way only pathways originating from the ancestors are considered.

However not all pathways leading back to ancestors would be known since some herds contributing to the present bull cohorts may have entered the recording system since 1964. Therefore three searching strategies were considered.

Strategy 1. All pathways were followed as far as possible.

Strategy 2. Relationships were sought through sires and bull dams, thus pathways did not contribute to a relationship if they passed through two females in succession. This required a modification to the getcoeff routine with a recursive parameter which when the dam of a dam was encountered instructed the routine to treat the dam's dam as unknown.

Strategy 3. Relationships were sought only through sires.

Table 2 *Proportions of genes in progeny derived from male and female parents of different ages when generation intervals are 6, 6·5 and 7 years*

Generation interval	Sex of progeny	Sex of parent	Age in years									
			2	3	4	5	6	7	8	9	10	11
7	M	M						0·2	0·2	0·2	0·2	0·2
		F					0·1	0·2	0·2	0·2	0·2	0·1
	F	M	0·125	0·25	0·125			0·1	0·1	0·1	0·1	0·1
		F		0·25	0·25	0·25	0·25					
6·5	M	M					0·1	0·2	0·2	0·2	0·2	0·1
		F				0·2	0·2	0·2	0·2	0·2		
	F	M	0·125	0·25	0·125			0·1	0·1	0·1	0·1	0·1
		F		0·25	0·25	0·25	0·25					
6	M	M					0·2	0·2	0·2	0·2	0·2	
		F			0·2	0·2	0·2	0·2	0·2			
	F	M	0·125	0·25	0·125		0·1	0·1	0·1	0·1	0·1	
		F		0·25	0·25	0·25	0·25					

The potential advantages of strategies 2 and 3 were that the pathways included in the tracking were more likely to have complete information. The disadvantage, however, was that fewer pathways were tracked. In strategy 3, only one path per descendant was tracked in any year. In strategy 2, the number of pathways leading from the descendant that were tracked increased exponentially as the tree was traversed upwards. This exponential rate of growth was, however, smaller than if the full tree had been traversed, thus each year it was expected to lose a further proportion of the relationships. Gene flow methods (Hill, 1974) were used to calculate the expected contribution from each search strategy. The gene flow matrices were constructed using the parameters given in Table 2. Three different matrices were used to model the decreasing generation interval from L=7 to 6 years. The expected contribution of strategy 1 was calculated assuming that starting in 1957, the breeding scheme operated for 7 years with L=7, then 13 years with L=6·5, and from then onward with L=6 years. For strategy 2, the coefficients in the gene flow matrix from female parents to female progeny were set to zero and for strategy 3, the coefficients from all female parents were set to zero. These expected contributions were then used to assess the completeness of the pedigree searching. Lindström (1969 and 1978) showed extensive variation in age among male and female parents and this was modelled in the gene flow matrices (cf. Tables 1 and 2).

The observed contributions were then used to estimate rates of inbreeding in the Finnish Ayrshire population. The estimates were subject to various assumptions which include: (i) missing pathways traced back to ancestors *pro rata* with known pathways; (ii) contributions to cohorts in 1986, 1987 and 1988 were close to convergence, and (iii) generation intervals are known in order to convert rates per generation to rates per year. The adequacy

Table 3 *Numbers of bulls (N) included in the data base according to year of birth*

Group of 'ancestors'								Group of 'descendents'	
Year	N	Year	N	Year	N	Year	N	Year	N
1958	37	1965	55	1972	221	1979	269	1986	272
1959	25	1966	90	1973	299	1980	316	1987	205
1960	38	1967	79	1974	315	1981	281	1988	230
1961	28	1968	111	1975	296	1982	267		
1962	35	1969	121	1976	294	1983	220		
1963	30	1970	139	1977	305	1984	291		
1964	26	1971	171	1978	281	1985	293		

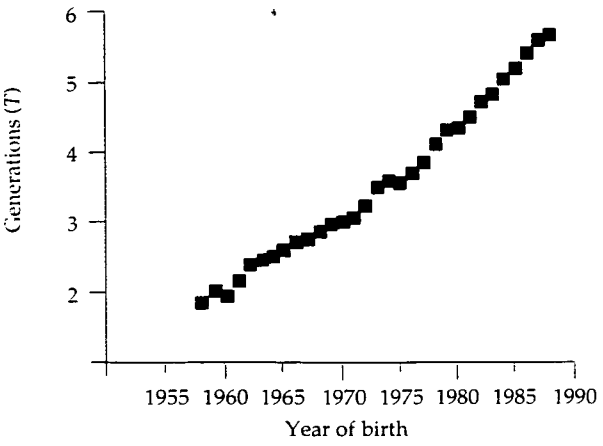


Figure 1 The number of generations known (T) for bulls according to year of birth.

of these assumptions and the effect of violations will be addressed in the **Discussion.**

Results

Numbers of bulls

Table 3 shows the numbers of bulls found in the data base according to year. The numbers for the years 1986 to 1988 exclude bulls which are sons of imported sires. A total of 219 sires were included in the ancestral generation.

Pedigree information and coefficients of inbreeding and relationship

Figure 1 shows the amount of pedigree information (T) present in units that are 'discrete generation equivalents' (see **Methods**). Analysis of the data shows non-linearity ($P < 0.001$) with slopes of 0·095 in 1965 and 0·153 in 1981. Although the reciprocal of the slope is not identical to the generation interval there is an association and the change in slope is consistent with the reduction in generation interval (Table 1).

Figure 2 shows the inbreeding coefficient F for each bull cohort. Ignoring data from before 1974 the regression of F on year had a slope of 0·0016 per year. Progressively dropping 1974, 1975 and 1976 increased the estimate to 0·0018 (s.e. 0·0001) and dropping further points made little further change. The removal of initial points is justified in that with overlapping generations several years will be required before the rate of inbreeding stabilizes.

When calculated using all the available pedigree including ancestors born before 1958, the average coefficients of relationship rose steadily over the full period from 1958 to 1988 (Figure 3). The regression on year for cohorts born after 1963 was 0·0030 (s.e. 0·0002).

Contributions

Strategy 1. The results of the gene flow showed that the expected sum of the contributions over four generations was 0·513. Since generations are overlapping the value is not exactly 0·5 and also because of the change in generation interval the expected ultimate sum of the contributions to the ancestors (to descendants born in later years) will be 0·505. Results are expressed as proportions of 0·513.

Figure 4 shows the histogram of the magnitude of the contribution among ancestors. Of the initial 219

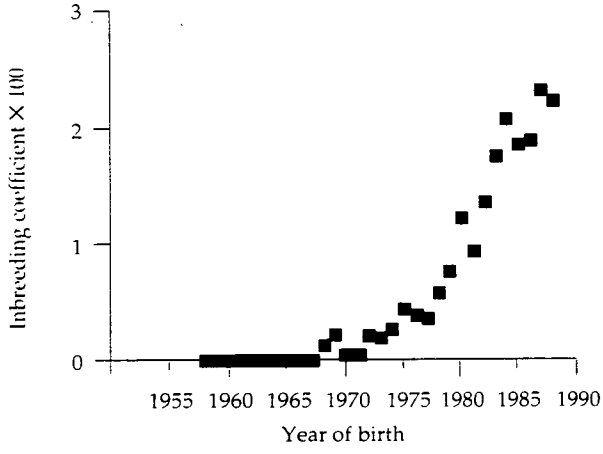


Figure 2 The mean inbreeding coefficient of bull cohorts according to year of birth.

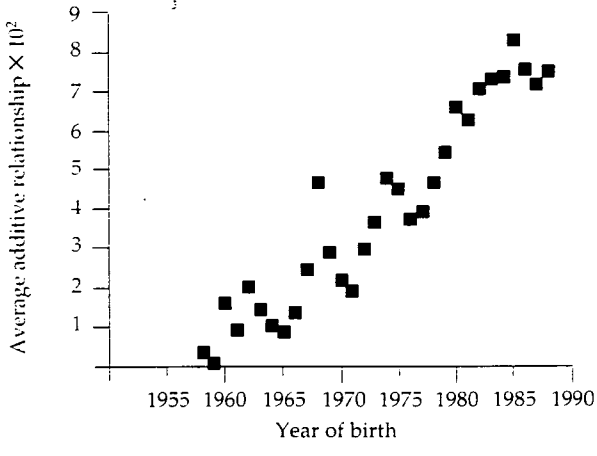


Figure 3 Average additive relationship within age groups among bulls born in years 1958-1990.

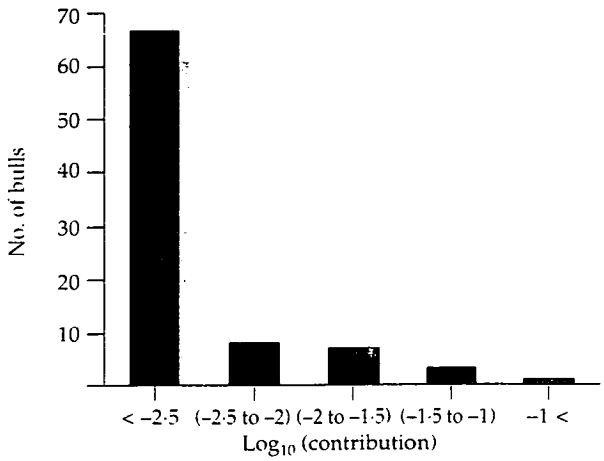


Figure 4 A histogram of proportional contributions after a full search among the 86 bulls with positive contributions.

ancestors, only 86 made a contribution of any size. The total of all the contributions was only 0·3, i.e. 58% of that expected. From Figure 4 it can be seen that the contributions among the 86 are highly skewed even on a log scale. There were 10 bulls contributing more than the expected value. These 10 bulls are listed in Table 4. An analysis was made of the 'missing' contributions and these were found to arise from three sources: (i) a large proportion of all pathways were traced back to A. Lier AAA 32605, an imported Norwegian bull born in 1968, whose contributions was 0·132; (ii) generation skipping was found where an important sire (Reima AAA 23597) was found to be born before the ancestors but whose influential sons were born after 1964, i.e. after the ancestors; (iii) missing pedigrees. Half the contribution of A. Lier AAA 32605 might be considered to belong to the ancestral generation (through the sire of A. Lier), however, even after accounting for items (i) and (ii) there is a shortfall of contributions totalling 0·08.

Table 4 *Proportional contributions of ancestor bulls having the largest impact on descendant groups in different search strategies (The bulls included are the top 10 from the full search)*

Bull name	Herd book no.	Contributions† Strategy 1	Strategy 2	Strategy 3
Mäki-Mattilan Inssi	AAA26350	0·25	0·30	0·18
Anttilan Mimro	AAA29107	0·07	0·09	0·19
Kreivilän Putte	AAA28200	0·05	0·07	0·21
Kunnarlan Kenraali	AAA27881	0·04	0·04	0·06
Lohko	AAA28899	0·02	0	0
Kumpuharjun Läiskä	AAA28049	0·02	0·01	0
Puotilan Akseli	AAA24844	0·02	0·02	0
Överby Rabbe	AAA26006	0·01	0·01	0
Syvärannan Maitoralli	AAA26705	0·01	0·02	0·01
Närin Erä	AAA25145	0·01	0·01	0

† In search strategy 1, the whole pedigree was tracked; in strategy 2, female to female paths were ignored; in strategy 3, only male to male paths were considered.

Strategy 2. The results of an analysis of gene flow showed that with the generation intervals assumed a total contribution of 0·313 from ancestors to descendants should be expected. The remainder of the 0·513 expected from the full search is lost by ignoring the grand-dams. This proportion of total contributions found using search strategy 2 would be expected to decine exponentially to zero over time.

Only 30 bulls make contributions but a strong correlation exists between the results of the two strategies (Table 4). One bull Forssan Lukko AAA 27850, not included in Table 4, had a proportional contribution 0·014 (cf. 0·01 in strategy 1), but no other bulls had a proportional contribution of 0·01. The total contributions observed were 0·197. However, if allowance is now made for the case of generation skipping in pedigrees and for the sire of A. Lier AAA 326505, the total contribution is 0·305, very close to the expected 0·313. Therefore it would appear that the pedigree searching of sire and maternal grandsire suffers little from loss of pedigree data.

Strategy 3. Gene flow analysis showed that the expected contributions of ancestors to descendants totalled 0·101. This proportion declines rapidly over years since only one pathway is tracked per descendant. Only nine bulls made contributions along this pathway and five of these are given in Table 4. The remaining four bulls, Forssan Lukko AAA 27850, Forssan Hakim AAA 25711, Ruikan Maileri AAA 28730 and Anttilan Jyrä AAA 27436 made contributions of 0·025, 0·013, 0·020 and 0·002, respectively. A total of 0·071 of the expected contribution was found. As with strategy 2 the impact of A. Lier AAA 326505 and a skipped generation explain the short fall. The ranking of contributions was changed from previous strategies. Analysis of expectations from gene flow suggested the contributions were sensitive to the particular cohort of ancestors and descendants, thus this searching strategy would be expected to be less robust.

Relationship of contributions to breeding values

The contributions were related to the estimates of breeding value of the ancestors for milk and protein

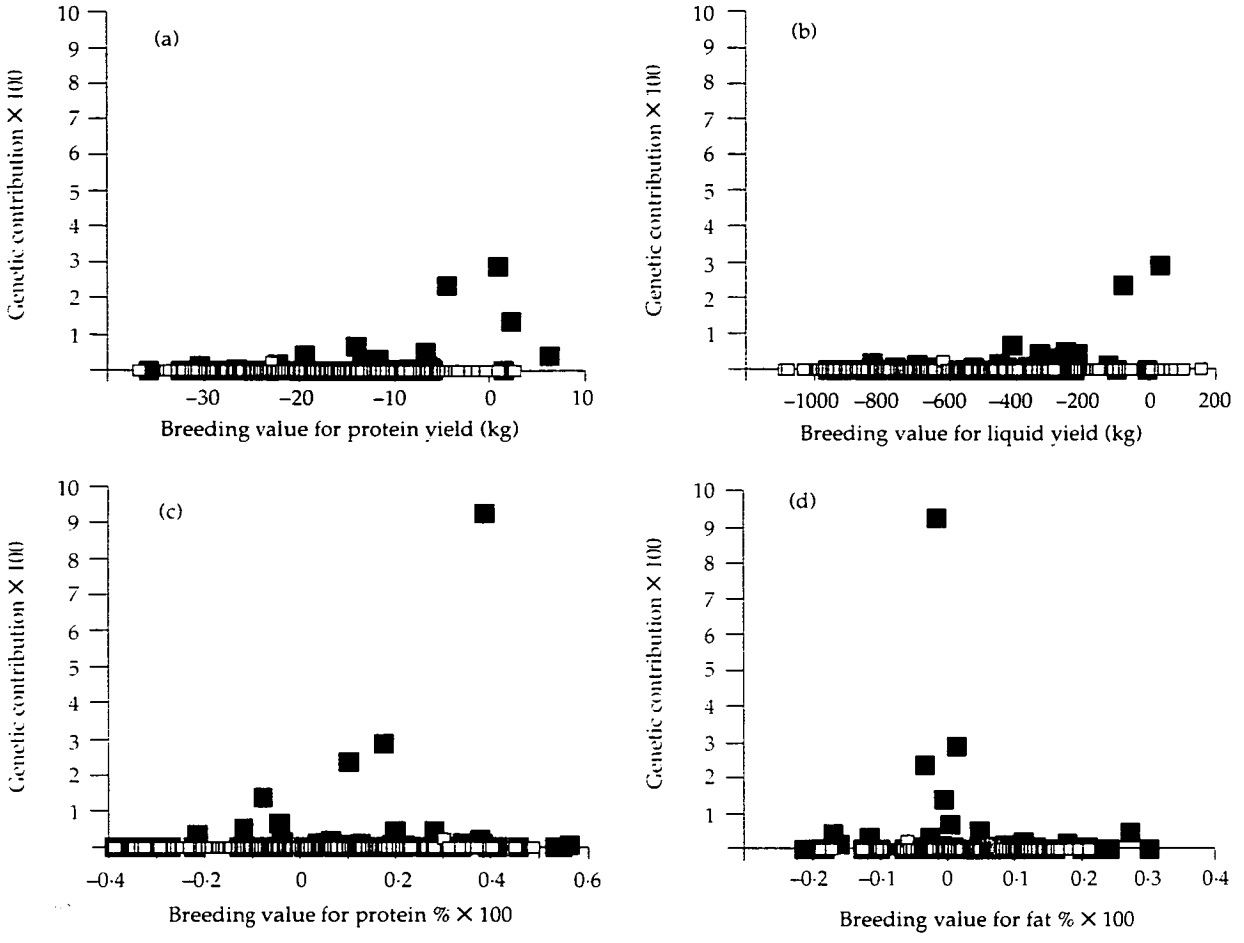


Figure 5 The relationship of genetic contributions obtained using strategy 2 with breeding values for (a) protein yield (kg); (b) milk yield (kg); (c) protein % (× 100); and (d) fat % (× 100). (□, squared accuracies < 0·75; ■, squared accuracies ≥ 0·75).

yield and fat and protein percentage. The best linear unbiased predictors of breeding values were obtained using the national data records. Unfortunately the data on the progeny test for the ancestors was not included so many of the bulls with low contributions were evaluated with low accuracy. Of the 219 bulls 128, 83 and 60 had squared accuracies greater than 0·5, 0·75 and 0·87 respectively. Figure 5 shows the relationship between the contributions obtained in strategy 2 and the breeding values.

Linear and quadratic regressions of the contributions on estimated breeding values were conducted separately for milk and protein yield and for fat and protein percentage. The analysis was carried out first for all ancestors (including those with a contribution of zero) and was repeated including only those ancestors with a squared accuracy of 0·75 or more. In all analyses the results are heavily influenced by the position of Mäki-Mattilan Inssi AAA 26350; however with and without restriction the contribution, measured by all three strategies increased with the estimated breeding value for protein yield and for milk yield ($P < 0·05$). A multiple linear regression upon all four estimated breeding values was also conducted; curiously, emphasizing the lack of robustness of these regressions, whilst the F-test using the 4 d.f. in the numerator was significant ($P < 0·05$) the single statistically significant coefficient was protein percentage. At best, with regression on all four breeding values including only those ancestors with squared accuracy of 0·75 or more the residual variance of the contributions was 0·877, 0·876 and 0·811 of the variance before regression for strategies 1, 2 and 3 respectively.

Prediction of inbreeding

Wray and Thompson (1990) showed that the rate of inbreeding was approximated by 0·25 times the sum of the squared contributions (as defined here) over all the M male and F female parents of a generation i.e. $\Delta F \propto \frac{1}{4}(\Sigma_{i=1}^{M} r_i^2 + \Sigma_{i=1}^{F} r_i^2)$ where r_i is the contribution of individual i. Where the observation of whether a pathway is missing or not is independent of the pedigree, the theory developed by Wray and Thompson (1990) predicts a missing pathway would trace back to the ancestors in the same relative proportions as the known pathways. We may therefore use the definition with this assumption to predict inbreeding rates in the Finnish Ayrshire population.

With these data the squared contributions from the males may be obtained directly. In the calculation we have: (i) assumed that the sire of A. Lier AAA 32605 was born in the period of the ancestors with a contribution of one half of that of Lier; and (ii) included Reima AAA 23597 who was born in 1957, before the ancestors, but was an influential sire whose important sons were born after the ancestral generation. Strategy 1 will be used to describe the calculation fully.

Using the results of strategy 1, by squaring the contributions obtained from the pedigree searching and summing over all ancestors, the sum of squared contributions for males was 0·02763. This omitted the missing pathways: in strategy 1 the sum of the identified contributions was 0·431 rather than 0·513. With the assumption that the missing pathways lead back to the ancestors in the same proportions as those that have been tracked then the estimate should be scaled by $(0{\cdot}513/0{\cdot}431)^2$. Furthermore over the long term it is expected from the analysis of gene flow that the total contribution of this generation will shrink from 0·513 to 0·505 and so the squared contributions require scaling by $(0{\cdot}505/0{\cdot}513)^2$. Finally, to estimate the male component of the inbreeding rate per generation the sum over male ancestors requires scaling by the ¼ giving a value of 0·0097 (i.e. 0·02763 × 1·413 × 0·969 × 0·25).

The sum of squared contributions from female ancestors can be inferred from the result for male ancestors. Firstly, it should be noted that F>>M, so important terms are to the order of M^{-1} (denoted $O(M^{-1})$) and those of $O(F^{-1})$ may be neglected. **Appendix 1** shows that to $O(M^{-1})$ the sum of squared contributions may be taken to be equal to the residual sum of squares after regressing the contribution of a male upon its breeding value. The regression is, however, conditional on the male being selected for breeding which differs from those presented in the previous section which included bulls irrespective of whether or not their contribution was zero. To obtain the correct residual sum of squares the regression needs to be repeated using only the 86 sires with positive contributions in the strategy 1 search; the minimum and mean squared accuracy of the estimated breeding values was 0·31 and 0·83 respectively. A residual sum of squares of 0·01536 was obtained after linear regression of breeding values on protein, protein percentage, fat percentage and yield. As for the male component of ΔF, this also requires scaling for missing pathways and ultimate contributions. This results in a value of 0·01536 × 1·413 × 0·969 × 0·25 = 0·0054 for the component of inbreeding per generation due to females. The resulting component was conservative in one sense since it was assumed that Reima AAA23597 and the sire of A. Lier AAA32605 would lie precisely on the regression line and so they made no addition to the residual. Combining the results for males and females gives ΔF = 0·0147 per generation. It is notable that in this case the squared errors were the dominant feature of the rate of inbreeding. To obtain annual rates of inbreeding the estimates require scaling by L^{-1}, where L is the generation interval. Although L is changing over time (see Table 1) a representative value for the period of 6·5 years has been used. This gives a value of ΔF = 0·00227 per year.

This process may be repeated for strategies 2 and 3 as if the additional information from other search strategies was unavailable. In both of these the female contribution was estimated using the residual from the regression of contributions on breeding values and including only those bulls for which positive contributions were found using the particular search strategy.

For strategy 2, the 30 ancestors with positive contributions had estimated breeding values with a minimum and mean squared accuracy of 0·51 and 0·93 respectively. The estimate of ΔF obtained was 0·0151 per generation, and 0·00232 per year.

For strategy 3, only nine ancestors could be included in the regression and these had estimated breeding values with a minimum and mean squared accuracy of 0·51 and 0·93 respectively. With the small number of ancestors, the residual sums of squares (for estimating female squared contributions) was taken from the best regression using only one variable and not the multiple regression described above for strategies 1 and 2. The estimate of ΔF obtained was 0·0125 per generation and 0·00193 per year. None of the regressions was in fact significant and if regressions were ignored the estimates of ΔF were increased to 0·0133 and 0·00205 respectively.

Discussion

Inbreeding in Finnish Ayrshires

There have been at least two previous studies of inbreeding in Finnish Ayrshires. Lindström and Maijala (1971) studied seven or more generations including ancestors dating back to 1894. They found the additive genetic relationship increased from 1915 to 1965 at a rate of 0·0011 per year and inbreeding at a rate of 0·0005 per year. They noted that there was evidence for the rate of inbreeding increasing to 0·0009 per year following the introduction of AI Mäntysaari and Stranden (1991) examined the inbreeding coefficient of AI sires born between 1960 and 1981 using an approximate method. With all available pedigree information the estimated inbreeding coefficients in 1981 was 0·0079.

The results of this study show that since the study of Lindström and Maijala (1971) coefficients of relationship and inbreeding are increasing faster than before. The rate of inbreeding according to identity by descent along the known pedigrees was 0·0018 per year, and the increase in the average coefficient of relationship was 0·0030 per year; both are more than double the previous estimates. This may be due to several factors: a decrease in generation intervals (Lindström, 1978) will certainly have contributed and possibly more intense selection through a greater awarenes of the use of AI. The estimate of 0·0018 per year is clearly a lower bound because of the missing pedigree information. The method used by Mäntysaari and Stranden (1991) underestimated the average inbreeding coefficients (as recognized in their study) since with full pedigree searching of known pedigrees, the average inbreeding coefficient of bulls had reached 0·0093 in 1981 (in this study 1981 was also lower than 1980 and 1982).

The rate of inbreeding per generation (0·0147 estimated from strategy 1) in the Finnish Ayrshire population exceeds the optimum for total economic merit (0·0125 per generation) suggested by Goddard and Smith (1990) even though when compared on an annual basis the rate is below the optimum. This is because of the difference in generation intervals assumed, 6·5 and 5 years respectively. The index used by Goddard and Smith was a simple economic weighting of progress and rate of inbreeding and since the breeding systems were the same i.e. progeny testing, the most appropriate comparison might be per generation (since in these circumstances the generation interval can be regarded as a scaling factor for both progress and inbreeding). Another approach towards defining maximum desirable rates of inbreeding was considered by Meuwissen and Woolliams (1995), which balanced inbreeding depression and gain in fitness through natural selection. This resulted in recommendations of the order of 0·006 to 0·007 per generation, much less than those observed in the Finnish Ayrshire. The maximum rates recommended were even lower if the selection index was negatively correlated with fitness, but in Finland indices were adjusted to remove such deleterious correlations when they are identified. As with Goddard and Smith (1990) the components of the criteria of Meuwissen and Woolliams (1995) are entirely genetic in nature and so comparisons are again perhaps most relevant per generation. Goddard (1992), although giving broadly similar recommendations to Goddard and Smith (1990) for Holsteins in optimum rates of inbreeding, introduces annual economic discounting of costs and benefits and this makes it less clear how rates of inbreeding may be compared when generation intervals differ.

The relationships of the breeding values of the Finnish Ayrshire ancestors with their contributions would suggest that selection among this group of ancestors has led to an increase in milk and protein yield with little progress in composition.

Prediction of inbreeding

The paper has examined the rates of inbreeding using contributions from an ancestral generation (Wray and Thompson, 1990) and the estimated annual rates were from 1·07 to 1·29-fold greater than those from identity by descent through known pedigrees with the estimates containing most information being the upper end of this range. The applicability of this approach to overlapping generations may be inferred from combining the proof of W. G. Hill in the thesis of Wray (1989), which relates contributions to ΔF via a drift argument, and that of Hill (1979) which develops the conceptual framework for considering overlapping generations.

The derivation relating contributions to inbreeding by Wray and Thompson (1990) was for individuals from the generation after an unselected base. Further analysis (J. A. Woolliams, unpublished) shows that the rate of inbreeding with selection may depend on the nature of base population, but the proof of Wray and Thompson remains valid. Therefore, the prediction of inbreeding given here is valid for a base population defined as the generation prior to the 'ancestral' generation.

Perhaps the most critical assumption made was that missing pathways trace back to ancestors *pro rata* with known pathways. If an alternative assumption was made, that the missing pathways were traced back to the ancestors known to have

contributed but with equal probability, then in strategy 2 the estimate would be barely one half of that observed from the known pedigree (results not shown). However, after several generations in a panmictic population the *pro rata* assumption will hold although variation observed in contributions among 1986, 1987, 1988 cohorts show that the contribution had yet to stabilize fully. The lack of convergence across cohorts would be expected to lead to an underestimate of the squared contributions, and hence ΔF, when selection indices are entirely based upon performance. This is because the variance in expected contribution among ancestors increases until convergence is reached (Wray and Thompson, 1990).

The estimates of annual rates of inbreeding are also dependent on the estimate of generation level. It should be noted that the rates per generation do not require this estimate. This problem is the reverse of that encountered when calculation is by regression of the observed coefficients of inbreeding on year, in that the estimate of generation interval is required to obtain rates per generation. Although the generation interval used was only approximate, variation of 0·5 years in the chosen value of 6·5 years would only lead to a proportional error of 0·08.

In the method used by Van Raden (1992) to overcome incomplete pedigrees there is an implicit assumption that an individual's unknown parent itself has parents which are drawn at random (i.e. with equal probability) from the parents of the previous generation. However, this assumption will not hold if descendants are known to exist since the pathways are more likely to trace back to superior ancestors. This becomes more serious the higher up the pedigree the missing information occurs. With this conservative assumption the estimated inbreeding rates increased some 1·14-fold when missing information was corrected for in the Ayrshire population of the United States of America studied by Van Raden (1992).

The estimate of strategies 2 and 3 provided an estimate of the rate of inbreeding which was calculated from as little as 0·3 of all pathways that go back to male parents. In the case of strategy 3, the estimate was produced in the absence of any female information in the pedigree. The methods described therefore allow estimation of ΔF with incomplete information. Notwithstanding the precision of the estimates of ΔF the measures using strategies 2 and 3 are simple to calculate and provide a straightforward means of evaluating the use of genetic variation within a breeding scheme.

Acknowledgements

Both authors are grateful to Finnish Animal Breeders Association (FABA) for allowing access to their data on the Finnish Ayrshire population and to Dr J. Juga and his colleagues at FABA for allowing us to use their estimate of the current generation interval in this population.

References

Hill, W. G. 1974. Prediction and evaluation of response to selection with overlapping generations. *Animal Production* **18:** 117-139.

Hill, W. G. 1979. A note on effective size with overlapping generations. *Genetics, USA* **92:** 317-322.

Goddard, M. E. 1992. Optimal effective population size for the global population of black and white dairy cattle. *Journal of Dairy Science* **75:** 2902-2911.

Goddard, M. E. and Smith, C. 1990. Optimum number of bull sires in dairy cattle breeding. *Journal of Dairy Science* **73:** 1113-1122.

Lindström, U. and Maijala, K. 1971. Development of the coefficients of inbreeding and relationship in the Finnish Ayrshire. *Zeitschrift für Tierzüchtung und Züchtungsbiologie* **87:** 335-347.

Lindström, U. 1969. Genetic change in milk yield and fat percentage in artificially bred populations of Finnish dairy cattle. Ph.D Thesis, University of Helsinki. *Acta Agralia Fennica* **114.**

Lindström, U. 1978. Selection intensity for milk yield in 1970-1977 in the Finnish Ayrshire. *Journal of the Scientific Agricultural Society of Finland* **50:** 445-454.

Miglior, F., Szkotnicki, B. and Burnside, E. B. 1992. Analysis of levels of inbreeding and inbreeding depression in Jersey cattle. *Journal of Dairy Science* **75:** 1112-1118.

Mäntysaari, E. A. and Stranden, I. 1991. Animal model evaluation for production and reproduction traits in Finnish dairy cattle. *Proceedings of the forty-second Annual Meeting of the European Association of Animal Production, Berlin.*

Meuwissen, T. H. E. and Woolliams, J. A. 1995. Required effective sizes for livestock populations. *Theoretical and Applied Genetics* In press.

Simonen, S. 1950. Ayrshireföreningens i Finland historica 1901-1951. In *History of Finnish Ayrshire Society.* Helsinki.

Van Raden, P. M. 1992. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. *Journal of Dairy Science* **75:** 3136-3144.

Woolliams, J. A. and Thompson, R. 1994. A theory of genetic contributions. *Proceedings of the fifth world congress on genetics applied to livestock production, Guelph* **19:** 127-134.

Woolliams, J. A., Wray, N. R. and Thompson, R. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genetical Research* **62:** 231-242.

Wray, N. R. 1989. Consequences of selection in finite populations with particular reference to closed nucleus herds of pigs. *Ph.D. thesis, Edinburgh University.*

Wray, N. R. and Thompson, R. 1990. Prediction of rates of inbreeding in selected populations. *Genetical Research* **55:** 41-54.

(*Received* 28 *February* 1995—*Accepted* 23 *March* 1995)

Appendix 1

In the following derivation emphasis is placed on the origin of terms that give rise to terms of $O(M^{-1})$ in ΔF. Wray and Thompson (1990) show $\Delta F = \frac{1}{4}(ME(r^2_{i(m)}) + FE(r^2_{j(f)}))$ where M is the number of male parents and F the number of female parents. Consider a hierarchical mating system with random mating and selection. Assume the selection has a form in which the correlation between different generations of selection arise through breeding values only; this occurs with mass selection or with sib-indices but not with evaluation using best linear unbiased prediction where, for example, phenotypic records (which will include components of environmental origin) will influence estimation of breeding values over many generations. Then consider an individual male $i(m)$ mated to $d(= FM^{-1})$ mates $j(f)$ $(j = 1,...,d)$. Then conditional on breeding values $A_{i(m)}$ and $A_{j(f)}$, $r_{j(f)} = \mu_j + \varepsilon_j$, where μ_j is a function of $A_{i(m)}$ and $A_{j(f)}$, namely $\mu_j = (2F)^{-1}(1 + \phi_m A_{i(m)} + \phi_f A_{j(f)})$ for some ϕ_m and ϕ_f (Woolliams and Thompson, 1994); $\text{cov}(\mu_j, \varepsilon_j) = 0$; and $\text{Var}(\varepsilon_j) = \sigma^2_j$ with $E[\sigma^2_j] = \sigma^2$ for all j. Therefore,

$$E[r^2_{j(f)} \mid A_{j(f)}, A_{i(m)}] = E[\mu^2_j] + \sigma^2_j = (4F)^{-2}(1 + \phi^2_m A^2_{i(m)} + \phi^2_f A^2_{j(f)}) + \sigma^2_j$$

and

$$FE[r^2_{j(f)}] = (4F)^{-1}(1 + \phi^2_m v_m + \phi^2_f v_f) + F\sigma^2$$

where $v_x = E[A^2_{i(x)}]$. Therefore the only terms of $O(M^{-1})$ in the component of ΔF due to females are those contained in $F\sigma^2$.

For $i(m)$, $r_{i(m)} = \sum_{j}^{d} r_{j(f)} = \sum_{j}^{d} \mu_j + \sum_{j}^{d} \varepsilon_j$. Conditional on the breeding values of $i(m)$ and mates $\Sigma\mu_j = (2M)^{-1}(1 + \phi_m A_{i(m)} + \phi_f d^{-1} \Sigma A_{j(f)})$. For small F the ε_j are negatively correlated since they are constrained to sum to zero but, for large F these may be regarded as independent and $E[\sum_{j}^{d} \varepsilon_j)^2] = d\sigma^2$.

Therefore $ME[r^2_{i(m)}] = (4M)^{-1}(1 + \phi^2_m v_m + \phi^2_f d^{-1} v_f) + F\sigma^2$.

Expanding terms $(4M)^{-1}d^{-1}\phi^2_f v_f$ is $O(F^{-1})$, which implies that the components of ΔF due to males that have $O(M^{-1})$ are $(4M)^{-1}(1 + \phi^2_m v_m)$ + terms of $O(M^{-1})$ in $F\sigma^2$. The expected sums of squares removed by regression of $r_{i(m)}$ on $A_{i(m)}$, where $i(m)$ is used for breeding is $(4M)^{-1}(1 + \phi^2_m v_m)$. This leads to the required result that the components of ΔF of $O(M^{-1})$ arising from female ancestors are equal to those in the residual sum of squares after regression of $r_{i(m)}$ on $A_{i(m)}$. This conclusion is in accord with the full evaluation of ΔF derived by Woolliams, Wray and Thompson (1993).

Paper 32

Expected genetic contributions and their impact on gene flow and genetic gain

by

J. A. Woolliams[1], P. Bijma[2], B. Villanueva[3]

[1] Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, U.K.
[2] Wageningen Agricultural University, Wageningen, The Netherlands
[3] Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, U.K.

Keywords: gene flow, genetic gain, pedigree, genetic contributions, selective advantage
Tel: +131 527 4200
Fax: +131 440 0434
E-mail: john.woolliams@bbsrc.ac.uk

INTRODUCTION

Within selection theory very little attention has been given to the relationship between the proliferation and diminution of individual genetic contributions within a population and properties of the population such as genetic gain. This is perhaps surprising since the development of the pedigree over generations provides the framework for the passage of genes through the population, forming the link between our understanding of individual genotypes and the way such genotypes influence the population. Such an understanding provides answers to, for example: the relative importance of individuals within a generation; where genetic change has arisen; how quickly the change generated has spread through the population; with what precision are we able to predict this change; how is genetic change related to the loss of variation; and how does genetic change in one generation relate to that in a subsequent generation. These questions have no general framework within which they can be answered although some special cases have been investigated (e.g. Villanueva *et al.*, 1996; Villanueva and Woolliams, 1997).

This study has the objective of deriving the expectations of pedigree development. In doing so it will develop the pioneering work of Wray and Thompson (1990), and more latterly the studies of Woolliams *et al.* (1993; mass selection), Wray *et al.* (1994; sib indices) and Woolliams and Thompson (1994). Firstly, the concept of genetic contributions will be considered in relation to genetic gain, and a general formula for gain will be proved. This expression of gain was given by Woolliams and Thompson (1994) but without detailed proof. The expected contribution of individuals over a number of generations will be derived, and the relationship of the long-term contribution to gain will be used to show the consistency between the theory developed on contributions and classical theory (e.g. Bulmer, 1983). Extensions to the theory of expected contributions will be given for overlapping generations and selection using best linear unbiased predictors (BLUP), and for genetic inheritance that is more complex than

direct and additive. Finally, the concept of the generation interval will be re-evaluated as a natural extension of the contribution theory. Many of the detailed results will be derived assuming an equilibrium. The use of the formulae developed will be shown in examples applied to discrete generations (sib indices, BLUP selection, and imprinted variation) and overlapping generations.

MATERIALS AND METHODS

Definitions and basic notation: The concept of genetic contributions was introduced by James and McBride (1958) and was developed by Wray and Thompson (1990). Given the fundamental nature of the concept to the paper, it will be re-stated. The genetic contribution of an ancestor i born at time t_1 to an individual j born at time t_2 ($>t_1$), is the proportion of the genes of j that derive by descent from ancestor i. This is different from the definition used by Wray and Thompson (1990) who multiplied this proportion by X_m+X_f (where X_m and X_f are the number of male and female parents in a generation), but as shown by Woolliams *et al.* (1993) a contribution is more usefully defined without this re-scaling. It is also distinct from the numerator genetic relationship which considers shared genes, and not those restricted to descent: thus full-sibs make no genetic contribution to each other although they have a genetic relationship >0. However, Wray and Thompson (1990) make clear the intimate relationship between these two concepts in their decomposition of the numerator relationship matrix.

The notation will be defined to allow extensions to overlapping generations. Therefore contributions will not be defined only in terms of individual ancestors within sex, but more generally as individuals within categories, where the categories define both age and sex and potentially breeding use (e.g. nucleus females and other females). Therefore over a lifetime an individual will move through various categories. An initial objective is to show the relationship between contributions and rate of gain, and for this there is no need to identify details of the

category of an individual and what is happening to the categories over time. However, for the concept of gene flow developed later, the tracking of categories is required. Therefore in order to keep notation minimal at any given stage, the notation for contributions will be developed through the paper, and a balance between consistency and simplicity has been attempted.

The following notation will be used initially: $r_{i,t_1}(j,t_2)$ is the contribution of ancestor i that was *born* at time t_1 to individual j born at time t_2; and $r_{i,t_1}(t_2)$ will define the mean over all the newborn cohort at time t_2 (½ of the mean for newborn males plus ½ of the mean for newborn females). Further reduced forms are for the long-term contributions of i, $r_{i,t_1} = r_{i,t_1}(t)$ as $t \to \infty$; t_1 in this context is generally not required and r_i is used. T_m males and T_f females are scored in each cohort, at random or systematically within families, and only scored individuals are candidates for breeding opportunities.

Populations will be assumed to be panmictic, which will mean that the contribution an individual ancestor makes to a population will tend to a constant for all individuals born in later cohorts, i.e. for each i $\mathrm{Var}[r_{i,t_1}(j,t_1+k)] \to 0$ as $k \to \infty$ (Wray and Thompson, 1990). This constant is r_i and will depend not only on the lifetime breeding use of i, but also upon its breeding value and other selective advantages both genetic and non-genetic, and chance factors.

In discrete generations, let $\boldsymbol{Z}_{t_1}$ be the matrix describing the passage of genes from cohort t_1 to cohort $t_1 + 1$, $\boldsymbol{r}_{t_1}$ the vector of long-term contributions for individuals in the cohort born at t_1, $\mathbb{1} = (1,1...1)^T$; then if $\boldsymbol{Z}_{t_1 ji} = ½$ depending if i is a parent of j and 0 otherwise, $\mathbb{1}\boldsymbol{r}_{t_1}^T = \mathit{limit}\ (k\to\infty)\ \boldsymbol{Z}_{(t_1+k)}\ldots\boldsymbol{Z}_{t_1+2}\boldsymbol{Z}_{t_1+1}\boldsymbol{Z}_{t_1}$. More generally, for overlapping generations, where $\mathbb{A}_t^*$ is the augmented numerator relationship matrix for all individuals born up to and including cohort t (Grundy *et al.*, 1998) then $\mathbb{1}\boldsymbol{r}_{t_1}^T = \mathbb{A}^*_{t(t,t_1)}$ as $t \to \infty$, i.e. the sub-matrix defined by the rows of cohort t and columns of cohort t_1.

Rates of gain: Decomposition of the augmented numerator relationship matrix $\mathbb{A}^*$ allows the re-

expression of the breeding value of an individual j born at time t as a sum of terms involving contributions from all other individuals born up to and including time t:

$$A_{j,t} = \sum_{k=1}^{t} \sum_{i} r_{i,k}(j,t)\, a_{i,k} + \sum_{i^*} r_{i^*,0}(j,t)\, A_{i^*,0}$$

where $a_{i,k}$ is the Mendelian sampling term of i born at time k, which is that part of its breeding value independent of its parental breeding values (say A_s and A_d) i.e. $a_{i,k} = A_{i,k} - \frac{1}{2}A_s - \frac{1}{2}A_d$. The second term with i^* is to allow for the base population, not necessarily unselected, where it is assumed that parents are unknown and so all the genetic information prior to $t = 0$ is vested in this base information. Define G_t, the genetic merit at time t, by $G_t = \frac{1}{2} \sum_{j\ males} A_{j,t} + \frac{1}{2} \sum_{j\ females} A_{j,t}$ then $G_t = \sum_{k=1}^{t} \sum_{i} r_{i,k}(t)\, a_{i,k} + \sum_{i^*} r_{i^*,0}(t)\, A_{i^*,0}$. The contribution of i to the cohort of its birth, $r_{i,t}(t)$ is taken to be $(2T_x)^{-1}$ since each sex will comprise half the breeding genes, and prior to selection this is shared equally among the candidates of sex x. Since $E[a_{i,k}] = 0$, the cross-products ra are related to the covariance between r and a, thus sustained genetic gain is related to the generation of covariance between contributions and Mendelian sampling terms.

Let the gain made by selection in cohort t be $\Delta G_t = G_{t+1} - G_t$, and $\Delta r_{i,k}(t) = r_{i,k}(t+1) - r_{i,k}(t)$ then:

$$\Delta G_t = \sum_{k=1}^{t} \sum_{i} \Delta r_{i,k}(t)\, a_{i,k} + \sum_{i^*} \Delta r_{i^*,0}(t)\, A_{i^*,0} \qquad (1)$$

Since the population is panmictic the covariance of the contributions and the Mendelian sampling terms must tend to a constant and so $\Delta r_{i,k}(t)\, a_{i,k} \to 0$ as $t \to \infty$ for a fixed k. In particular the terms for the base population terms in equation (1) tend to 0; and (identifying males and females separately)

$$E[\Delta G_t] = \sum_{k=1}^{t} \sum_{x=m,f} T_x E[\Delta r_{i(x),k}(t)\, a_{i(x),k}] \qquad (2)$$

If an equilibrium is approached (as will be the case with the infinitesimal model), the expected *change* in covariance between r and a will depend only on t-k and not on k *per se* i.e. only on the time elapsed since birth and not on the time of birth:

$E[\Delta r_{i(x),k}(t)\, a_{i(x),k}] \rightarrow E[\Delta r_{i(x),k+j}(t+j)\, a_{i(x),k+j}]$.

Therefore:

$$\sum_{k=1}^{t} E[\Delta r_{i(x),k}(t)\, a_{i(x),k}] = \sum_{k=0}^{t-1} E[\Delta r_{i(x),t}(t+k)\, a_{i(x),t}]$$

and for large t in an equilibrium this sum approaches $E[r_{i(x),k}\, a_{i(x),k}]$ for any k.

Therefore for a sufficiently large t, $E[\Delta G_t] = E[\Delta G_{eq}]$ and substitution of these results into equation (2) gives:

$$E[\Delta G_{eq}] = T_m E[r_{i(m)}\, a_{i(m)}] + T_f E[r_{i(f)}\, a_{i(f)}] \qquad (3)$$

or alternatively, $E[\Delta G_{eq}] = T_m \mathrm{cov}(r_{i(m)}, a_{i(m)}) + T_f \mathrm{cov}(r_{i(f)}, a_{i(f)})$.

An equivalent expression to equation (1) can be given as a continuous function of time, and at time t:

$$E[\Delta G(t)] = \sum_{x=m,f} T_x \int_0^t D(t,x,u)\, du$$

where $D(t,x,u) = d/dt\ E[r_{i(x),u}(t)\, a_{i(x),u}]$. For an equilibrium using the infinitesimal model, $r_{i(x),u}(t)$ is a function of $z = (t-u)$; $d/dt\ E[r_{i(x),u}(t)\, a_{i(x),u}] = d/dz\ E[r_{i(x),u}(z)\, a_{i(x),u}]$, with $r_{i(x),u}(z)$ independent of u, and so the subscript u can be neglected. Substituting z for u in the integration gives equation (3).

Expected long-term contributions

Framework for general solution: To exploit the relationships between the contributions and rates of gain and rates of inbreeding (Wray and Thompson, 1990) it is necessary to replace r, which is an observed variable, with predictions. To develop this aspect it will be necessary to modify slightly the notation used: in particular, it will be necessary for breeding categories to be explicit; so $i(q)$ will be an ancestor in category q. However, with the assumption of equilibrium it will only be necessary to consider the parents from all categories (maybe of different ages) at a single time point (which will be taken to be t=0; but this does not imply it is an unselected base population). Thus $r_{i(q)}(j(p),t)$ denotes the contribution of the parent i in category q at time 0 to

j in category p at time t (the essential difference with the previous notation is that *t no longer refers to the time of birth*).

The expected long-term contribution of individual $i(q)$ will be defined conditional on a vector of variables s, that influence the number of its own offspring that are selected (e.g. an expected breeding value (EBV) from BLUP) and (or) which may influence the selection of subsequent descendants (e.g. a prediction error for the EBV) i.e. $\mu_{i(q)}=E[r_{i(q)}] =\alpha_q + \beta_q^T s$. Initially s will be assumed to be a single variable, e.g. the breeding value (A) of an ancestor which was used for mass and sib-index selection by Woolliams *et al.* (1993) and Wray *et al.* (1994). The expected lifetime long-term contribution of an individual i will be the sum of the expected long-term contributions for all categories that i belonged to over its lifetime.

The objective of the following is to define a set of achievable steps which can be followed to derive expected contributions even in complex breeding schemes. These steps are: (i) (for overlapping generations only) to determine the gene flow to selected individuals in the current period from parents in previous periods; (ii) to develop a regression model for the expected number of offspring a parent may have (with coefficients λ, forming a matrix Λ); and (iii) to develop a second regression model describing the relationship of the selective advantages of a selected offspring with those of the parent (with coefficients π, forming a matrix Π). The details of the regression models are given in Appendices 1 and 2.

Equilibrium will be used to obtain concise formulae for the expected long-term contributions. The use of equilibrium can be illustrated best by returning to an analogy used by Wray and Thompson (1990). Suppose a random number, n, of random variables u occur subsequent to, and independent of, the decision on n. Let $v = \sum u$, then $E[v] = \mu_n \mu_u$. If the expectations of v and n are linearly related to a variable $A-\bar{A}$ (i.e. expressed as a deviation from its mean) i.e. $E[v] = \alpha + \beta(A-\bar{A})$, $E[n] = 1 + \lambda(A-\bar{A})$; and that the variables u have the same

distributional property as v, that is $E[u] = \alpha + \beta(A^* - \overline{A^*})$, but their indexing variable A^* is different from A and has a regression relationship with A given by $(A^* - \overline{A^*}) = \pi(A - \overline{A})$ plus a random error. Then, equating terms in A resulting from $E[v] = \mu_n \mu_u$, gives a relationship between α and β, namely:

$$\beta = (1 - \pi)^{-1} \lambda \alpha \tag{4}$$

To make the analogy more direct, note that the long-term contribution of individual i is given by:

$$r_i = \tfrac{1}{2}(\sum r_{j(m)} + \sum r_{j(f)}) \tag{5}$$

where the sums are taken over its male and female offspring; since unselected offspring have no long-term contribution these sums may be restricted to the selected offspring. Using mass selection as a model (where the long-term contribution can be described as a regression on breeding value; Woolliams *et al.*, 1993), if we are in an equilibrium then the relationship of the long-term contribution of i with its own breeding value, expressed as a deviation from the mean of those selected from the cohort $A - \overline{A}$, is identical in form to the relationship of the long-term contribution of the selected offspring to their own breeding value deviation $(A^* - \overline{A^*})$. Furthermore the breeding value of the selected offspring can be predicted from a linear regression on the parental breeding value. Therefore the analogy is between v and the long-term contribution of the parent, u and the long-term contribution of a selected offspring, and between n and the numbers of male or female offspring selected.

The remaining tasks are: (i) to generalize the result to multiple categories, which will include discrete generations where there are just two categories (males and females) and to develop the regression equations required; and, (ii) to extend the result for several variables describing selective advantage (i.e. s rather than simply A alone).

Extension to multiple categories: The method can be extended to multiple categories (n_c)

which cover sex, age and breeding purpose. The expected long-term contributions for individual i in category q, $\mu_{i(q)}$, can be represented by $\alpha_q + \beta_q (A_{i(q)} - \overline{A_q})$, where $(A_{i(q)} - \overline{A_q})$ has expected value zero in the *selected* individuals which make up category q. The vectors α and β refer to the vector of coefficients for all categories and are both of dimension (n_c x 1).

The concept of gene flow (Hill, 1974) is used, but the development of Hill does not account for the inheritance of selective advantage which is critical to the development of genetic contributions. A further consequence of this selective advantage is that the origin of the parents of a selected group will be dependent on the degree of selection taking place. As groups grow older the selection pressures change, so the origin of the parents of the selected individuals will also change with age.

Step 1, defining the gene flow matrix $\boldsymbol{G}$: Denote the standard gene flow matrix for a breeding structure (Hill, 1974) by $\boldsymbol{G}_0$. The key elements in the matrix are $g_{0,pq}$ representing the proportions of genes in the newborn cohort from which category p will be selected at some point in the future, that arise from category q in the previous time period. To obtain the equilibrium long-term contributions a modified matrix is required ($\boldsymbol{G}$, of dimension n_c x n_c) in which each row represents a category of *selected* individuals (rather than newborn), and with the elements g_{pq} of each row representing the transfer of genes through breeding from the parents in the different categories q. With discrete generations and the standard two pathways, $\boldsymbol{G}$ = (½,½ | ½,½) always.

Step 2, defining Λ: A regression model is required to relate the selection score in category p (i.e. number of offspring selected to breed in that category) to the breeding value of a parent in category q. In principle, the regression model will depend on both parent and offspring category, and will form a matrix of coefficients of dimension n_c x n_c. With random selection the regression model is $2\boldsymbol{NGN}^{-1}$; where $\boldsymbol{N}$ is a diagonal matrix with elements equal to the number of selected

individuals in category q (i.e. X_q), and the 2 corrects for the rows of matrix $\boldsymbol{G}$ summing over ages to ½ for each sex. To maintain the gene flow form, the regression model will be denoted by $2\boldsymbol{N}(\boldsymbol{G}\otimes(\boldsymbol{I}+\Lambda[A])\boldsymbol{N}^{-1}$, where $[A]$ is used as a convenient notation to tag the regression coefficients for selective advantage. The form may be rationalized: $\boldsymbol{N}^{-1}$ converts a selected ancestor in category q into a proportion of the gene pool of origin, the $2\boldsymbol{G}$ and Λ predict the proportions in the offspring gene pools in relation to their origin, and the $\boldsymbol{N}$ converts the proportions back into a number of descendants.

Step 3, defining Π: A second regression model is required for the regression of the breeding value of the *selected* offspring on the breeding value of the parent. In principle these, too, depend on both the category of offspring and parent, giving a matrix Π, with π_{pq} representing the coefficient for offspring category p and parent category q.

Step 4, solutions: Using equation (5), the assumption of equilibrium, and collecting terms in the constant term and those for the regression coefficient separately gives:

$$\alpha_q = \sum_p X_q^{-1} g_{pq} X_p \alpha_p \tag{6a}$$

$$\beta_q[A] = \sum_p X_q^{-1} g_{pq} X_p \lambda_{pq} \alpha_p[A] + \sum_p X_q^{-1} g_{pq} X_p \beta_p[A^*] \tag{6b}$$

Substituting $\pi_{pq}[A]$ for $[A^*]$ and collecting terms:

$$(\boldsymbol{N}\alpha) = \boldsymbol{G}^T(\boldsymbol{N}\alpha) \tag{7a}$$

$$(\boldsymbol{N}\beta) = (\boldsymbol{I}-\boldsymbol{G}^T\otimes\Pi^T)^{-1}(\boldsymbol{G}^T\otimes\Lambda^T)(\boldsymbol{N}\alpha) \tag{7b}$$

where $\otimes$ denotes element-by-element multiplication of the matrices. Therefore $\boldsymbol{N}\alpha$ is a right eigenvector of $\boldsymbol{G}^T$ with eigenvalue 1 (it has such an eigenvalue since all rows of $\boldsymbol{G}$ sum to 1) with elements that sum to L^{-1}, where L is a generation interval. The occurrence of the transposes of the matrices in the formulae arises from defining the matrices in the form customarily used for gene flow (Hill, 1974), i.e. rows defining progeny categories and columns defining parental categories. For discrete generations with the standard two pathways, $\alpha = (½X_m^{-1}, ½X_f^{-1})$ always.

Extension to multiple variables (s): With multiple variables conferring selective advantage, $\mu_{i(q)} = \alpha_q + \beta_q^T(s_{i(q)} - \overline{s}_q)$. Let n_s be the number of variables in s, then each category has a sub-ordering within it of the n_s variables. α remains a vector of length n_c but β is a vector of length $n_c n_s$. The matrix Λ is order n_c x $n_c n_s$, and Π is $n_c n_s$ x $n_c n_s$. Then if $p(v)$ in the following summation represents variable v in category p:

$$\beta_{q(u)}[s_u] = \sum_p X_q^{-1} g_{pq} X_p \lambda_{pq(u)} \alpha_p[s_u] + \sum_p X_q^{-1} g_{pq} X_p \sum_v \beta_{p(v)} \pi_{p(v)q(u)}[s_u] \qquad (8)$$

Equations (7) are again obtained with: (i) the definition of $\otimes$ being extended to mean the multiplication of the sub-matrix defined by category p progeny and category q parents i.e. π_{pq} by the element g_{pq}; and (ii) in (7b), $N\beta$ is replaced by $N'\beta$ where N' is now diagonal of dimension $n_c n_s$ x $n_c n_s$ with each X_q repeated n_s times.

A further refinement of α: A further improvement can be made in the estimation of α, which corresponds to a second order approximation. The g_{pq} account for the different selective advantages among the categories of the parents at the time of selection but the advantages or disadvantages are inherited in part by the offspring and $\alpha_q = \sum_p X_q^{-1} g_{pq} X_p (\alpha_p + \beta_p^T \boldsymbol{d}_{pq})$, where $\boldsymbol{d}_{pq} = E[s_p \mid \textit{category } q \textit{ parent}] - \overline{s}_p$. In schemes such as mass and index selection the deviations $\boldsymbol{d}_{pq}$ are simply derived from the parental differences. Therefore, after rearranging terms in equations (6a) and (6b), and where $\boldsymbol{D}$ is dimension ($n_c n_s$ x n_c), with sub-matrix pq equal to $\boldsymbol{d}_{pq}$:

$$(N\alpha) = (G^T + (G^T \otimes D^T)(I - G^T \otimes \Pi^T)^{-1}(G^T \otimes \Lambda^T))(N\alpha) \qquad (9)$$

Whilst α is still defined as a right eigenvector, the matrix is more complex. When generations are discrete and with the standard two pathway model $\boldsymbol{D} = 0$.

Development of contributions over time: To simplify the notation the development of the finite time contributions is given for the single selective advantage A. For category q, a *selected* individual at time 0, has a vector (dimension n_c x 1) of contributions to *selected* individuals in

all categories at time t given by $\boldsymbol{c}_q(t) + \boldsymbol{b}_q(t)\ (A_{i(q)} - \overline{A}_q)$. Then $\boldsymbol{c}_q(0)$ is zero except for X_q^{-1} in the qth position and $\boldsymbol{b}_q(0) = 0$. A further vector of regressions is required $\boldsymbol{f}_q(t)$ which denotes the pooled regression of selective advantage for the categories at time t on the selective advantage of the individual at time 0. By definition $\boldsymbol{f}_q(0) = 1$.

It is critical to note that the contributions at time t to the selected individuals in category p of age *age(p)*, will depend on the consequences of the selection upon the parental gene pool at time *t-age(p)*. Therefore the complete spectrum of contributions to the categories at time t will depend not only on the previous state but on previous states up to the maximum age of the parents in the breeding scheme. As a result gene flow with selection is a process that needs to be defined in terms of the state variables of the preceding n time periods, where n is the maximum age of any parental group. Define $\boldsymbol{G}_p$ to be the n_c x n_c matrix consisting of zeros except for the single row corresponding to category p. Then

$$\begin{aligned} \boldsymbol{c}_q(t) &= \sum_p \boldsymbol{G}_p\, \boldsymbol{c}_q(t-age(p)) \\ \boldsymbol{b}_q(t) &= \sum_p \boldsymbol{G}_p\, \boldsymbol{b}_q(t-age(p)) + \sum_p X_q^{-1}(\boldsymbol{G}_p \otimes \Lambda)\, \boldsymbol{f}_q(t-age(p)) \\ \boldsymbol{f}_q(t) &= \sum_p (\boldsymbol{G}_p \otimes \Pi)\, \boldsymbol{f}_q(t-age(p)) \end{aligned} \tag{10}$$

As $t \to \infty$, $\boldsymbol{c}_q(t) \to \boldsymbol{G}^t \boldsymbol{c}_q(0)$, $\boldsymbol{b}_q(t) = \boldsymbol{G}^t(\boldsymbol{G} \otimes \Lambda)\ (\boldsymbol{I} - \boldsymbol{G} \otimes \Pi)^{-1}\, \boldsymbol{f}_q(0) X_q^{-1}$, and $\boldsymbol{f}_q(t) \to 0$. The convergence of $\boldsymbol{f}$ to 0 reflects the diminishing effect of ancestors over time on the selection advantage of their descendants. The panmictic assumption ensures that both $\boldsymbol{c}_q(t)$ and $\boldsymbol{b}_q(t)$ converge to a vector with all elements equal, namely α_q and β_q respectively. Therefore, combining these results over categories by ordering them into columns, the following is obtained:

$$\mathbb{1}\alpha^T = \mathbb{G}^t \mathbb{N}^{-1}$$

$$\mathbb{1}\beta^T = \mathbb{G}^t(\mathbb{G} \otimes \Lambda)\ (\mathbb{I} - \mathbb{G} \otimes \Pi)^{-1}\, \mathbb{N}^{-1}$$

and this pair leads back to equations (7). (Where $\mathbb{1}=(1, \ldots, 1)^T$).

The discrete time contributions with the refinement in estimating α is given in Appendix

3. An example of application is given in the results.

Expected long-term contributions and rates of gain: For any one individual i the total long-term contribution is the sum of its long term contributions as it moves through the different categories over its life time i.e. $r_i = \sum_q r_{i(q)}$. Define $S_{i(q)} = 1$ if i is selected in category q, 0 otherwise, then:

$$E[r_i \mid S_{i(q)},\ q=1,...n_c] = \sum_q S_{i(q)} E[r_{i(q)} \mid S_{i(q)} = 1]$$

Thus the estimates of the expected long-term contribution will depend on the information available on $S_{i(q)}$. If the information on $S_{i(q)}$ is incomplete, $S_{i(q)}$ may also need prediction and

$$E[r_i \mid Information] = \sum_q E[S_{i(q)} \mid Information]\ E[r_{i(q)} \mid S_{i(q)}=1]$$

When the expected long-term contribution is expressed in terms of all the components of the breeding value, in particular the Mendelian sampling term, the expected long-term contribution is sufficient for the prediction of genetic gain. If an individual is unselected then $r=0$ and furthermore the sum of $S_{i(q)}$ over all candidates is X_q; and the rate of gain can be expressed solely in terms of the selected individuals. This yields:

$$E[\Delta G_{eq}] = \sum_q X_q E[\mu_{i(q)} a_{i(q)}] \qquad (11)$$

where now the expectations are conditional on being selected as a parent rather than unconditional as was the case in equation (3).

If $\mu_{i(q)} = \alpha_q + \beta_q^T (s_{i(q)} - \overline{s}_q)$ then equation (11) immediately decomposes the gain into two components: the first, $\sum_q X_q \alpha_q E[a_{i(q)}]$, is the gain obtained from selection within families, which is accrued at the time of selection of the ancestor; whilst the second, $\sum_q X_q \beta_q^T E[(s_{i(q)} - \overline{s}_q) a_{i(q)}]$, is defined only in terms of the selected individuals within the categories, and is the gain obtained from selecting between the ancestors once selected, and so represents the between family gain. Since the between family gain is explicitly defined in terms of the selective advantages, the gain can be decomposed into the components arising from each

of the advantages separately.

The covariance between the Mendelian sampling term and s following the selection of the ancestor can be calculated using standard index theory.

Derivation of λ: The approach used in the examples to derive Λ has followed the method of Robertson and its development by Wray and Thompson (1990). When selection is completely determined by the index, the regressions may be calculated as the product of the regression of the selection score of the offspring (S=1 if selected, S=0 otherwise) on the index, the regression of the offspring index on the vector of variables conferring selective advantage (s) and the number of offspring per parent. The last term may also be defined in terms of s, but this is not developed here. The general approach adopted is outlined in Appendix 1.

Derivation of Π: Appendix 2 gives a general derivation for Π, and shows that the sub-matrix of regression coefficients for s in the selected *progeny* of category p on the values of s in their parents in category q are approximated by $\pi_{pq} = (V_{pq} - k_p \sigma_I^{-2} \nu_p \nu_q^T) V_{qq}^{-1}$ where V_{qq} is the variance/covariance matrix for category q before selection in category p, V_{pq} is the covariances of the s_p and s_q, and k_p is the variance reduction parameter for selection in category p. This approach was used in all the results given below.

APPLICATION OF MODELS AND RESULTS

Expected long-term contributions for general sib indices in discrete generations: The rates of inbreeding for general sib-index of the form $I = b_1(P - \overline{P}_F) + b_2(\overline{P}_F - \overline{P}_H) + b_3\overline{P}_H$ was studied by Wray *et al.* (1993; WWT). Mass selection is a special case with $b_1=b_2=b_3=1$. WWT obtained results for expected long-term contributions, but their derivation was less direct than the method here. s will be the breeding value (A) alone, with other forms of environmental influences that are often considered (e.g. litter effects) omitted for simplicity. For discrete generations there are just two categories, males and females. It is assumed that the phenotypic variance σ_P^2=1, and

the unselected additive genetic variance is h_0^2.

The regression models required are derived from Appendices 1 and 2: $\lambda_{pq} = i_p\tau_q(2\sigma_I)^{-1}$ and $\pi_{pq} = \frac{1}{2}(1-k_p\tau_q\rho\sigma_a\sigma_I^{-1})$ where $\tau_m = b_3$ and $\tau_f = b_2(1-X_mX_f^{-1}) + b_3X_mX_f^{-1}$ and $\tau = 2(\tau_m + \tau_f)$. The τ_q values are twice the regression of the index on the breeding value of the parent of sex q.

Appendix 4 shows the calculation of β_m and β_f, with the results

$$\beta_q = \tfrac{1}{4}i(\tau+\tau_q)(\sigma_I+\kappa z)^{-1}X_q^{-1}$$

where $\kappa = [k\tau + \frac{1}{8}(\tau_m-\tau_f)(k_m-k_f)]$ and $z = \rho\sigma_A$. This form is nearly equivalent to that given by WWT, but their derivation proceeded on different (and more complex) lines: three points of difference should be noted: (i) WWT does not include the small $\frac{1}{8}(\tau_m-\tau_f)(k_m-k_f)$ term in κ; (ii) the indices of WWT are also explicitly scaled to $\rho\sigma_A\sigma_I^{-1} = 1$, but scaling does not change the ratio $\tau_q(\sigma_I)^{-1}$; (iii) in this paper it is explicit that predictions in equilibrium are being sought using equilibrium values of parameters such as σ_I.

Rates of gain from sib indices: The decomposition of the rates of gain is achieved using equation (11) and standard index theory. Within family gain for sex q is $\frac{1}{4}h_0^2i_q\tau_w\sigma_I^{-1}$ obtained from $\alpha_q = (2X_q)^{-1}$ and $E[a_{i(q)}|\ i\ selected] = \frac{1}{2}h_0^2i_q\tau_w\sigma_I^{-1}$ where τ_w, is the regression of the index I on $a_{i(q)}$:

$$\tau_w = b_1(1-n_F^{-1}) + b_2(n_F^{-1}-n_H^{-1}) + b_3n_H^{-1}$$

where n_F and n_H records contribute to the full- and half-sib means respectively (including the candidate).

Since $cov(a_{i(q)},A_{i(q)}) = \frac{1}{2}h_0^2(1-k_q\tau_wz\sigma_I^{-1})$ the between family gain arising from selection in sex q is $\frac{1}{8}h_0^2i(\tau+\tau_q)(1-k_q\tau_wz\sigma_I^{-1})(\sigma_I+\kappa z)^{-1}$. The difference between the sexes in the between family gain is generated from the first generation of progeny. After this first generation the contributions proliferate indirectly through progeny of both sexes used as parents and this generates the term common to both sexes that is described by τ.

The total gain, summed over both sexes, including both between and within families is:

$$\Delta G_{eq} = \tfrac{1}{2}h_0^2 i(\tau_w + \tau)(\sigma_I + \kappa z)^{-1} \qquad (12)$$

This uses the result $k_m\tau_m + k_f\tau_f = \tfrac{1}{2}(k_m+k_f)(\tau_m+\tau_f) + \tfrac{1}{2}(k_m-k_f)(\tau_m-\tau_f) = 2k\tau + \tfrac{1}{2}(k_m-k_f)(\tau_m-\tau_f)$.

Consistency with previous formulations: The expression for equilibrium ΔG_{eq} can be equated to the standard formula $\Delta G = i\rho\sigma_A = iz$. Equating the two forms result in a quadratic equation in z:

$$\kappa z^2 + \sigma_I z - \tfrac{1}{2}h_0^2[\tau_w + \tau] = 0 \qquad (13)$$

A check on the consistency between the methods given here with classical index theory for discrete generations can be made by noting that the quadratic equation in z is obtainable directly from the equilibrium conditions using standard index theory (result not shown):

$\sigma_A^2 = \tfrac{1}{2}h_0^2 + \tfrac{1}{4}\sigma_A(1-k_m\rho^2) + \tfrac{1}{4}\sigma_A^2(1-k_m\rho^2)$ and that $\mathrm{cov}(A,I) = \rho\sigma_A\sigma_I$. Thus neglecting the second order correction for the Bulmer effect in Appendix 2 appears to be implicit in standard index theory.

The quadratic equation (14) can be used to give reasonable estimates of equilibrium gain (iz) for indices when even using initial parameters. However, the estimates will not be precise since they assume σ_I constant, as well as $\Delta F = 0$; but unlike when $i\rho\sigma_A$ is calculated from initial parameters, they partially account for selection through the regressions π_{pq} and provide reasonable estimates even when using base parameters. Further improvements using (13) would require an iterative scheme in combination with:

$$\sigma_I^2 = \sigma_I^2 - \tfrac{1}{4}z^2([b_2^2(1-X_mX_f^{-1}) + b_3^2X_mX_f^{-1}](k+k_f) + b_3^2(k+k_m))$$

Table 1 shows estimates of gain obtained from various formulae using initial parameters and equilibrium gain achieved after several generations (with no inbreeding). With initial parameters iz consistently over-predicted and consistently had the largest errors whilst equations (12) and (13) underestimated gain; equation (13) was always closest. Table 2 partitions the

equilibrium gain achieved between and within families for each sex from applying equation (13). As the index weight places relatively more weight on pedigree the within family gains decrease. Even in the index where no weight is placed on the paternal half-sib means, more gain was obtained from the males, largely through the greater selection intensity during the selection within families. The index coefficients used are chosen for illustration only.

Long-term contributions for Best Linear Unbiased Predictors (BLUP): The analysis of individual long-term contributions can be extended to BLUP evaluation and indices based directly upon it. An initial consideration for the model proposed, is the form that μ_i will take. In sib-indices μ_i was determined by A_i since it is the only means by which a parent may influence its offspring over multiple generations. In BLUP this is no longer the case since the parental information contributes to the evaluation of the offspring, and so a parent's phenotypic record and its associated environmental influences continue to influence selection of offspring. Therefore the development of the estimated breeding value through time influences the long-term contribution.

However, there are logical guides to what should, and should not, be considered for the proposed model. It is clear that factors influencing the selection of offspring are valid and feasible predictors of the long-term contribution. This does not summarize the entire long-term contribution since the remaining prediction error will also have a residual effect influencing the selection success of subsequent descendants. The information available is summarized in the BLUP estimate $\hat{A}$ (the EBV) at any given time. Here three terms are considered for individual i in category q: $\hat{A}_{i(q)}$, the 'initial EBV' at the point of selection of the parent; $\delta\hat{A}_{i(q)}$, the 'increment' in the EBV at the point of selection of its progeny; and $\hat{e}_{i(q)}$, the remaining 'prediction error' of the parent at the selection of offspring. Therefore, $\mu_{i(q)} = \alpha_q + \beta_q^T(s_{i(q)} - \overline{s_q})$ where s_i is a vector comprising $\hat{A}_i$, $\delta\hat{A}_i$ and $\hat{e}_i$. Appendix 5 gives the derivation of Λ and Π necessary for the

determination of $\mu_{i(x)}$.

An example of the application is given in Table 3 where predictions are compared to simulations with selection based upon pseudo-BLUP as described by Wray and Hill (1989). The simulated hierarchical scheme is for 20 male parents, 40 female parents in discrete generations with 8 full-sibs per family. Heritability was 0.4. The $\boldsymbol{G}\otimes\Pi$ and $\boldsymbol{G}\otimes\Lambda$ matrices are given in the Appendix 5. Excellent agreement is found between simulations and predictions, both for the regressions and the total gain, even when base generation parameters are used.

The results show that the primary source of between family selection among ancestors in BLUP is the increment in the EBV between its own selection and that of its offspring. The initial EBV plays the least important role, with slightly more between family gain (in this example) coming from the prediction error.

Extensions to other inheritance modes in the absence of allelic interactions: The extensions of the model to other inheritance modes can be considered by defining the variables in s and their impact on λ_{pq} and π_{pq}. As an example, the situation with maternal imprinted variation is developed. For maternal imprinting, the breeding value can be split into A^+ inherited from the female parent and expressed, and A^- inherited from the male parent and not expressed. Define $\mathbf{s} = (A^-, A^+)$, with discrete generations giving two categories, m for males and f for females. In this case, λ_{pm} will be zero since the genes passed by the male do not influence selection of its offspring. However λ_{pf} will depend on both breeding values, since although A^- is not expressed in the female it is expressed in its offspring. For π_{pq}, there will be a dependence on both breeding values: genes passed by the male only affect A^-, and genes passed by the female only affect A^+; since genes passed by the male are not expressed the regression of parent on offspring is unaffected by selection.

$$G = \quad (\tfrac{1}{2}, \tfrac{1}{2} \mid \tfrac{1}{2}, \tfrac{1}{2})$$

$$\Lambda = \quad (0.0, 0.0, \tfrac{1}{2} i_m \sigma_P^{-1}, \tfrac{1}{2} i_m \sigma_P^{-1} \mid 0.0, 0.0, \tfrac{1}{2} i_f \sigma_P^{-1}, \tfrac{1}{2} i_f \sigma_P^{-1})$$

$$\Pi = \quad (\tfrac{1}{2}, \tfrac{1}{2}, 0.0, 0.0 \mid 0.0, 0.0, \tfrac{1}{2}(1- k_m h^2), \tfrac{1}{2}(1- k_m h^2)$$

$$\mid \tfrac{1}{2}, \tfrac{1}{2}, 0.0, 0.0 \mid 0.0, 0.0, \tfrac{1}{2}(1- k_f h^2), \tfrac{1}{2}(1- k_f h^2))$$

where $h^2 = \mathrm{Var}(A^+)/ \sigma_P^2$, and the phenotypic variance, σ_P^2, is the sum of the variance of A^+ and the environmental variance.

Predictions were made using base generation parameters and also when the equilibrium variance parameters were iterated to reach equilibrium. From these, predictions of β were made using equations (7). To calculate ΔG, the expected mean of the Mendelian sampling terms for selected individuals and the covariance with s for selected individuals were calculated using standard index theory. In the long-term, since this is imprinted variation, half the genes from an ancestor will be expressed in females, and half will be latent in males. Therefore predicted gains should be halved. Predictions were compared with simulations obtained for 20 male and 40 female parents with three full-sib offspring of each sex per female.

Table 4 shows very close agreement between simulations and predictions, including the prediction based upon the base generation heritability and phenotypic variance. The gains within families shown in the Table entirely arise through the expressed breeding value of the candidates (they achieved at the selection of the parents). Approximately 0.6 of the between family gains shown in the Table arise from selection on the latent breeding value of the candidates, and since the regressions are identical this effect may be ascribed to the larger genetic variance associated with this term (it is not reduced through the initial selection of the ancestors). The regression of genetic contribution on breeding values was greater in females than in males despite their greater number, which is not surprising given the mode of inheritance.

Overlapping generations: An example of application in overlapping generations is presented

here for mass selection with a fixed generational structure in a two-path scheme (i.e. there was no subdivision of breeding individuals into males to breed males, males to breed females etc.). The general approach is explained in more detail by Bijma and Woolliams (1998). The steps will be illustrated using a scheme with three categories: 20 males breeding each year at one year of age, 20 females breeding at one year of age and 20 females breeding at three years of age respectively. The number of offspring per litter is eight and the heritability was 0.4. The age groups not used for parents will be omitted, to give 3 x 3 matrices: males age one, females age one, females age three.

1. g_0, which denotes the genetic make-up of the newborn is (0.5, 0.25, 0.25). From g_0, and the number of parents and the family sizes, the intensities of selection (i_p) and variance reduction coefficients (k_p) were calculated for each category: i_p=1.647, k_p=0.817.

2. An initial ΔG was assumed was assumed as a starting point for iteration. In the following ΔG from gene flow using Hill (1974) after iterating to an equilibrium will be used: ΔG= 0.412.

3. The genetic value of the selected parents within male categories and within female categories was calculated as $i_p h^2 \sigma_P$ - ($age(p)$-1)ΔG; and deviations from the means within male and female categories were calculated δ = (0, +0.412,-0.412).

4. The genetic variance in the offspring were calculated using the pooled variance within categories plus between categories plus the Mendelian sampling variance:

$$\tfrac{1}{2} h_0^2 + \sum_p \left(\tfrac{1}{4}\sigma_A^2 2g_{0,p}(1-k_p h^2) + \tfrac{1}{4}(2g_{0,p})\delta_p^2\right)$$

This was 0.370, and the phenotypic variance σ_P^2 = 0.970.

5. G was calculated by an optimal truncation algorithm applied separately across male and across female categories: for sex x, the frequency of offspring from category p of sex x was given by $2g_{0,p}$ the mean of their offspring was $\tfrac{1}{2}\delta_p$, and the phenotypic variance among offspring within categories was $\sigma_P^2 - \sum \tfrac{1}{4}(2g_{0,p})\delta_p^2$ (where the sum removes the component of genetic

variance between categories of the same sex as p): in the first iteration each row of $\boldsymbol{G}$ was (0.5, 0.336, 0.164).

6. Π and Λ matrices were constructed according to the Appendices 1 and 2 respectively. For mass selection $\pi_{pq} = 0.5(1- k_p h^2)$, and $\lambda_{pq=}0.5 i_p \sigma_P^{-1}$. In the first iteration, $\Pi = 0.344\ \mathbf{1}\mathbf{1}^T$ where $\mathbf{1}^T = (1,1,1)$, $\Lambda = 0.836\ \mathbf{1}\mathbf{1}^T$, and $\boldsymbol{D} = \mathbf{1}$ (0,0.092,-0.188).

7. α and β were calculated according to equations (7b) and (9). In the first iteration, $(N\alpha)^T$= (0.395,0.289,0.106) and $(N\beta)^T$=(0.503, 0.338, 0.165).

8. The covariance of the Mendelian sampling term with the breeding values were calculated and ΔG was updated using equation (11); this uses the result that $E[a_{i(q)}] = \frac{1}{2} h_o^2 i_q \sigma_P^{-1}$ and after selection $cov(a_{i(q)}, A_{i(q)}) = \frac{1}{2} h_o^2 (1- k_q h^2)$.

9. Steps 3 to 8 were repeated to convergence.

Results from the iterations were $\alpha = (0.0200, 0.0149, 0.0050)^T$ and $\beta = (0.0255, 0.0171, 0.0084)^T$. The gain predicted within families was (0.134, 0.100, 0.034) for each category, and between families was (0.067,0.045,0.022), giving a total gain of 0.402. At equilibrium $\boldsymbol{G}$ was $\boldsymbol{1}$ (0.500,0.335,0.165). This was compared to simulation results for 1000 replicates: $\alpha = (0.0197, 0.0145, 0.0052)^T$ with a maximum s.e. of 0.0009; $\beta = (0.0249, 0.0175, 0.0071)^T$ with a maximum s.e. of 0.0004; and a total gain of 0.398 (s.e. 0.001). Thus very close agreement between simulations and predictions were obtained. As in discrete generations the gain from mass selection was evenly divided between males and females. The gene flow predicted using Hill(1974) is $\alpha = (0.0167, 0.0083, 0.0083)$; Hill(1974) makes no prediction of β.

The generation interval, defined by the time taken to turnover the genes once, was predicted from $(\Sigma X_q \alpha_q)^{-1}$ to be 1.25 (cf. 1.26 in the simulations) which was notably shorter than in the average age of the parents. This is because of the cumulative effect of the selective advantage of the younger age group of females; whilst they produced equal numbers of offspring

they produced more than twice as many parents. However, the generation interval was not predictable from the equilibrium G alone (i.e. accounting for a single generation of selective advantage) since this would have predicted an interval of 1.33 (i.e. 0.5×1 + 0.335×1 + 3×0.165).

To obtain the time course of the contributions Appendix 3 was used. This uses equations (8) which need the following matrices based on G:

G_1 (0.500, 0.335, 0.165 | 0.0, 0.0, 0.0 | 0.0, 0.0, 0.0)

G_2 (0.0, 0.0, 0.0 | 0.500, 0.335, 0.165 | 0.0, 0.0, 0.0)

G_3 (0.0, 0.0, 0.0 | 0.0, 0.0, 0.0 | 0.500, 0.335, 0.165)

The results are shown in Table 5, for the time course of contributions from category 2. The contributions converged in cohort 10.

DISCUSSION

This study has developed a framework for predicting the expected genetic contribution under a wide range of selection and inheritance models. This framework allows selection to be more properly accounted in comparison with existing gene-flow methods for overlapping generations and multiple breeding groups (e.g. Hill, 1974), and furthermore extends the methods to consider the differential gene flow among individuals within categories, an extension not hitherto achieved except in some special cases. The framework has been constructed by firstly modelling the selection process and the transfer of selective advantages within a single generation of selection, and secondly, extending this to multiple generations. Two regression models are required, both of which are derived using standard index theory: firstly, a model describing the expected number of selected offspring a parent may have (Λ); and the second describing the relationship of the selective advantages of a selected offspring with those of the parent (Π).

The framework has been developed to describe the expected genetic contribution over

all time horizons from the short-term to the long-term. The novel, closed formulae (7b and 9) produced for the expected long-term contribution of an ancestor rely on the assumption of equilibrium in the selection process. However this assumption is not necessary for the use of equations (10) where contributions are predicted over finite time periods, but clearly more effort may be required to define the changes in the parameters that are used to build the models. In the examples given for discrete generations, assuming infinite alleles generating the variation, the predictions for the regressions of contributions on selective advantages within a cohort showed considerable robustness to the use of initial parameters rather than equilibrium parameters. This may be ascribed to the Π-model accounting for the impact of selection in the parent on the selective advantages of the offspring, and therefore anticipating the changes in parameters that selection will bring. For non-equilibrium states the result will depend on the relative degree of departure in relation to the timescale of convergence of the contributions (approximately five generations).

In the development of the framework, the effects of inbreeding on parameters and progress have been neglected, but this is not a serious problem. Firstly, the timescale for the convergence of contributions is relatively small in comparison to the timescale for the effects of inbreeding on parameters in breeding schemes, especially where inbreeding is controlled to be at reasonable levels. The impact of individuals within a cohort is very largely decided within five generations, and even within this period, the scope for controlling an individual's contribution declines exponentially. A second reason is that schemes will most usefully be compared at the same rates of inbreeding, and so the neglect of inbreeding will be less likely to bias the comparisons made.

The expected long-term contribution has been described in a general linear form $\alpha_q + \beta_q^T(s_{i(q)} - \overline{s_q})$, where s is a vector of selective advantages for an ancestor i. Quadratic terms

in s have not been considered since their inclusion does not effect the prediction of rates of inbreeding unless higher moments than the variances of the components of s are considered (Wray and Thompson, 1994), and, judged by the results in this study, their omission does not lead to serious errors.

The α represent the proportion of genes that derive from the various categories as a whole and these can differ qualitatively from predictions using Hill (1974) since the earlier study does not account for the inheritance of selective advantages. In any newborn cohort (which may be sub-divided), even when generations overlap, the cohort of males is expected to have an equal long-term contribution to the cohort of females. However the division of that total contribution will vary according to breeding use, and this division has been accurately predicted with the current model. The framework presented here and that of Hill(1974) give the same prediction of α when selection is at random, since (i) elements of $\boldsymbol{G}$ are identical to those of $\boldsymbol{g}_0$, (ii) $\Pi = 0$, and (iii) $\Lambda = 0$.

The genetic contribution of an individual represents the expected impact its Mendelian sampling term has on the population, and so the sum of the total contributions from any one cohort is a natural measure of the rate at which genes in the population are renewed. In particular the rate measured by the $\Sigma\alpha$ places an emphasis upon those contributions that are destined to remain in the population in the long-term. Thus $(\Sigma\alpha)^{-1}$ is a measure of the generation interval, L. The generation interval defined by the long-term contributions is shorter than the traditional average age of the parents for the examples considered, because the younger breeding groups had a selective advantage and the progeny of older parents were less likely to be selected.

The need for a modified generation interval arising from the inheritance of the selective advantage has been considered previously (Rendel and Robertson, 1950; Bichard *et al.*, 1973; James, 1977), but this is the first time a comprehensive solution to the problem has been given.

It has long been accepted that parents whose purpose was to produce a commercial cohort outside the breeding population should not be included in the calculation of a generation interval. Furthermore, James (1977) moved the concept further by considering the generation interval calculated from only those parents with selected offspring. However the results show that this step is insufficient to account for the selective advantages over many generations. The average age of the parents and the definition of James (1977) may be viewed as a one-generation estimate of the generation interval and an iteration beyond this respectively. These concepts are examined in more detail by Bijma and Woolliams (1998). The average age of the parent will remain of operational significance in breeding schemes, but the genetic generation interval is more appropriately defined by the long-term contributions.

The consistency of the framework with other approaches for estimating gain in discrete generations is important, but this consistency does not extend to overlapping generations. The main approach for prediction of gain in overlapping generations is that of Rendel and Robertson (1959) which was given support from an apparent consistency with the gene flow approach of Hill (1974). However, this study shows that this consistency is not in fact justified. The estimates of equilibrium gain using contributions and Rendel and Robertson differ slightly from each other. The estimate of gain from contributions arises from the prospective analysis of the impact of a single cohort to the future population over the long-term. In contrast, the estimate of gain from Rendel and Robertson (1950) arises from a retrospective analysis of the impact of selection in the whole population to a single cohort. One technical reason why differences between these approaches will appear are the distributions assumptions made concerning selection differentials, since the overlapping structure increases the number of disparate truncated normal distributions being mixed together in a single cohort.

The second component of the expected long-term contribution is the linear regression on

the selective advantages of an individual (β) which describes the expected development of the pedigree within a category and, consequently, the expected extent of between-family selection that will occur as a result of the selection process. The between-family selection is responsible for the greater rates of inbreeding that can occur when selection is practised, and the control of the magnitude of the regression coefficients (and the components of $\boldsymbol{s}$) is an important aspect of methods to optimize the genetic gain with constrained inbreeding rates (e.g. Villanueva and Woolliams, 1997). The methods presented allow the between family gain to be decomposed into its constituent parts, as was done in the results.

The between-family selection may develop very quickly, so that its extent is largely established in the selection of the progeny, or more slowly. This time-course is controlled by $\boldsymbol{G}\otimes\Pi$ and powers of $\boldsymbol{G}\otimes\Pi$, which describe the decay of the ancestor's selective advantage through progeny (see equation for $\boldsymbol{f}(t)$ in (10)). In this context, the eigenvalues of $\boldsymbol{G}\otimes\Pi$ are of interest. In the example given for BLUP, the maximum eigenvalue of $\boldsymbol{G}\otimes\Pi$ was 0.18, which may be compared to 0.36 for mass selection with the same numbers of parents and initial heritability. Therefore it is clear that a higher proportion of the ultimate between-family selection generated by selection with BLUP is achieved in the first and second generations after the ancestor than is the case with mass selection. This difference has a consequence for the accuracy of the prediction of rates of inbreeding using techniques accounting for co-selection in one- and two-generations (Wray *et al.*, 1989), and explains why these methods are notably more accurate with BLUP selection than with mass selection (T.H.E. Meuwissen, personal communication).

The importance of predicting the development of genetic contribution is that aspects of risk in breeding schemes such as ΔF cannot be described without a knowledge of the dynamics of individual contributions. The importance of the expected genetic contribution is made greater by the result of Woolliams (1998) which showed that ΔF can be predicted from the expectation

alone. The framework presented here provides a step-by-step recipe for predicting this expected genetic contribuiton over multiple generations. In providing the results, particular approaches have been described to derive the necessary regression models (Appendices 1 and 2). However it is important to recognize that the details of these Appendices are not an integral part of the recipe and other approaches could replace them in the recipe to suit the needs of a particular study.

ACKNOWLEDGEMENTS

J.A. Woolliams gratefully acknowledges: the Ministry of Agriculture, Fisheries and Food (UK) for funding this work; the encouragement of Dr. P. Visscher and Dr. B. J. McGuirk; and Prof. A. Maki-Tanila and Prof. B. Kinghorn for providing opportunities to develop and complete it. PB gratefully acknowledges financial support from the Netherlands Technology Foundation (STW). And BV gratefully acknowledges financial support from the Biotechnology and Biological Sciences Research Council.

LITERATURE CITED

BICHARD, M., A. H. R. PEASE, P. H. SWALES and K. ÖZKÜTÜK, 1973Selection in a population with overlapping generations. Animal Production 17: 215-227.

BIJMA, P. and J. A. WOOLLIAMS, 1998 Prediction of genetic contributions and generation intervals in populations with overlapping generations under selection. Genetics (submitted).

HILL, W. G., 1974 Prediction and evaluation of response to selection with overlapping generations. Animal Production **18**: 117-139.

JAMES, J. W. 1977. A note on the selection differential and generation length when generations overlap. Animal Production **24**: 109-112.

JAMES, J. W. and G. McBRIDE, 1958. The spread of genes by natural and artificial selection in a closed poultry flock. Journal of Genetics **56**: 55-62.

RENDEL, J. M. and A. ROBERTSON, 1950 Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. Journal of Genetics **50**: 1-8.

VILLANUEVA, B. and J. A. WOOLLIAMS, 1997 Optimization of breeding programmes under index selection and constrained inbreeding. Genetical Research (Cambridge) **69**: 145-158.

VILLANUEVA, B., J.A. WOOLLIAMS and B. GJERDE, 1996. Optimum designs for breeding programmes under mass selection with an application in fish breeding. Animal Science 63: 563-576.

WOOLLIAMS, J. A. 1998 A recipe for the design of breeding schemes. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production 25: 427-430.

WOOLLIAMS, J. A. and R. THOMPSON, 1994 A theory of genetic contributions. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 19: 127-134.

WOOLLIAMS, J. A., N. R. WRAY and R. THOMPSON, 1993 Predictin of long-term contributions and inbreeding in populations undergoing mass selection. Genetical Research (Cambridge) **62**: 231-242.

WRAY, N. R. and W. G. HILL, 1989 Asymptotic rates of response from index selection. Animal Production **49**: 217-227.

WRAY, N. R., J. A. WOOLLIAMS and R. THOMPSON, 1990 Methods for predicting rates of inbreeding in selection populations. Theoretical and Applied Genetics **80**: 503-512.

WRAY, N. R., J. A. WOOLLIAMS and R. THOMPSON, 1994 Prediction rates of inbreeding in populations undergoing index selection. Theoretical and Applied Genetics **87**: 878-892.

APPENDIX 1: A GENERAL APPROXIMATION TO λ_{pq}

The regression of selection score of the unselected candidates of category p on the index I is given by $p_p i_p/\sigma_I$ (Wray and Thompson,1990). For a parent of category q, the regression of the candidate index on **s** for all the parents of category p that are of the same sex as category q (denote by **w**) was derived by standard index theory appropriate to the inheritance model under consideration.

For each offspring of the parent from group q the probability of selection can then be approximated by $p_p(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\boldsymbol{s}-\bar{\boldsymbol{s}}))$. The expected number of offspring for a parent of category p is then $n_p p_p(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\boldsymbol{s}-\bar{\boldsymbol{s}}))$ where n_p is the number of candidates in category p per parent. $n_p p_p$ is equal to or $2g_{0,pq}X_pX_q^{-1}$, where g_0 the proportion of genes among the newborn category p that derive from category q.

Considering category q parents, they have an average selective advantage given by $\overline{s_q}$ so the expectation is $2g_{0,pq}X_pX_q^{-1}(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\boldsymbol{s}-\overline{\boldsymbol{s}_q}) + i\sigma_I^{-1}\boldsymbol{w}^T(\overline{\boldsymbol{s}_q}-\bar{\boldsymbol{s}}))$. For sufficiently small deviations this is approximately $2g_{0,pq}X_pX_q^{-1}(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\boldsymbol{s}-\overline{\boldsymbol{s}_q}))(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\overline{\boldsymbol{s}_q}-\bar{\boldsymbol{s}}))$ where the last term in the product may be viewed as the additional selective advantage of category q, and so $g_{0,pq}(1 + i\sigma_I^{-1}\boldsymbol{w}^T(\overline{\boldsymbol{s}_q}-\bar{\boldsymbol{s}})) \approx g_{pq}$ and $\lambda_{pq} \approx i\sigma_I^{-1}\boldsymbol{w}$.

APPENDIX 2: DERIVATION OF π_{pq}

Let s_q be the vector of deviations of explanatory variables from their mean for a parent in category q and s_p for progeny in category p and likewise I_p be the index deviations for selection of category p. Let $s = (s_q^T \mid s_p^T \mid I_p)$ have the partitioned (co)variance matrix

$$\begin{pmatrix} V_{qq} & V_{qp} & \nu_q \\ V_{qp}^T & V_{pp} & \nu_p \\ \nu_q^T & \nu_p^T & \sigma_I^2 \end{pmatrix}$$

Before selection on I_p, s_q and s_p are expressed as regressions on I_p:

$$s_p = \sigma_I^{-2} \nu_p \, I_p + \epsilon_p$$

$$s_q = \sigma_I^{-2} \nu_q \, I_p + \epsilon_q$$

Equating $E[s_q s_q^T]$ to V_{qq} gives $E[\epsilon_q \epsilon_q^T] = V_{qq} - \sigma_I^{-2} \nu_q \nu_q^T$ and, similarly $E[\epsilon_p \epsilon_q^T] = V_{pq} - \sigma_I^{-2} \nu_p \nu_q^T$. After selection, Normal distribution theory infers that the regression coefficients on I_p are unchanged. Therefore, after selection

$$V_{pq}^* = (V_{pq} - k_p \sigma_I^{-2} \nu_p \nu_q^T), \; V_{qq}^* = (V_{qq} - k_p \sigma_I^{-2} \nu_q \nu_q^T)$$

Let π_{pq} be the matrix of coefficients of s_p on s_q after selection, then $\pi_{pq} = V_{pq}^{*T} \, V_{qq}^{*-1}$.

In the applications described this will be approximated by $\pi_{pq} = V_{pq}^{*T} \, V_{qq}^{-1}$. This is for three reasons: (i) simpler forms; (ii) it coincides with preceding published theory on genetic contributions; and (iii) such an assumption is implicit in standard index theory.

As an example with more than a single variable consider mass selection with random mating, where the vector of selective advantages explicitly includes the breeding value of the mate as well as the individual. In this case $s_{i(q)}$ has 2 variables $(A_{i(q)}, A_{i(q')})$, where $A_{i(q)}$ is the breeding value of i and $A_{i(q')}$ is the breeding value of its mate, and define $s_{j(p)}$ similarly for the selected progeny $j(p)$. $V_{pq} = (\tfrac{1}{2}\sigma_A^2(1-k_q h^2), \tfrac{1}{2}\sigma_A^2(1-k_q h^2) \mid 0, 0)$, $\nu_p = (\sigma_A^2, 0)$, $\nu_q = (\tfrac{1}{2}\sigma_A^2(1-k_q h^2), \tfrac{1}{2}\sigma_A^2(1-k_q h^2))$, $V_{qq} = \mathrm{diag}(\sigma_A^2(1-k_q h^2), \sigma_A^2(1-k_q h^2))$, resulting in $\pi_{pq} = (\tfrac{1}{2}(1-k_p h^2), \tfrac{1}{2}(1-k_p h^2) \mid 0, 0)$.

These are results of Wray and Thompson (1990). In this example it was chosen to obtain a fuller description of the expected long-term contribution by explicitly including the mate; ignoring the mate is valid for considering genetic gain, providing the matrices are appropriately constructed e.g. if mating had been assortative rather than random the covariance between parent and offspring breeding value would need to account for the mate implicitly.

APPENDIX 3: CONTRIBUTIONS OVER FINITE TIME WHEN α IS ESTIMATED AS A RIGHT EIGENVECTOR OF $(\boldsymbol{G}^T + (\boldsymbol{G}^T \otimes \boldsymbol{D}^T)(\boldsymbol{I} - \boldsymbol{G}^T \otimes \Pi^T)^{-1}(\boldsymbol{G}^T \otimes \Lambda^T))$

For category q, a *selected* individual at time 0, has a vector of contributions to *selected* individuals in categories at time t given by $\boldsymbol{c}_q(t) + \boldsymbol{b}_q(t)\ [A]$; $[A]$ denotes that the $\boldsymbol{b}$ is a vector of regression coefficients on selective advantages for differential contributions from within category q. Adjustment of equations (8) will be done assuming a univariate selective advantage. The approach taken is to use a modified form of equation 5:

$$r_{i(q)}(t) = \tfrac{1}{2} \sum_{offspring\, j \in category\, p} r_{j(p)}(t - age(p))$$

Therefore the expected contribution after t cohorts is calculated by considering the expected contributions of progeny in category p, for t-$age(p)$ cohorts.

Firstly, $\boldsymbol{c}_q(t)$ and $\boldsymbol{b}_q(t)$ are calculated according to equations (8). Then the following iterative scheme is applied where $\boldsymbol{c}^*$ and $\boldsymbol{b}^*$ are the solutions from the previous iteration.

$$\boldsymbol{c}_q(t) = \sum_p X_q^{-1} X_p g_{pq} (\boldsymbol{c}_p^*(t - age(p)) + \boldsymbol{b}_p^*(t - age(p))\, \delta_{pq})$$

$$\boldsymbol{b}_q(t) = \sum_p X_q^{-1} X_p (g_{pq} \lambda_{pq}) \boldsymbol{c}_p^*(t - age(p)) + \sum_p X_q^{-1} X_p (g_{pq} \pi_{pq}) \boldsymbol{b}_p^*(t - age(p))$$

The number of iterations required is approximately equal to the time horizon of interest.

APPENDIX 4: THE EXPECTED LONG-TERM CONTRIBUTIONS FOR SIB INDICES IN DISCRETE GENERATIONS DESCRIBED BY WRAY *et al.* (1994)

For the sib-indices described by Wray *et al.* (1993): $G = (\frac{1}{2}, \frac{1}{2} \mid \frac{1}{2}, \frac{1}{2})$ and $\pi_{pq} = \frac{1}{2}(1 - k_p\, \tau_q\, z\, \sigma_I^{-1})$, $\lambda_{pq} = (2\sigma_I)^{-1} i_p \tau_q$, where $z=\rho\sigma_A$. From equation (7a) $\alpha = (\frac{1}{2}X_m^{-1}, \frac{1}{2}X_f^{-1})^T$.

Applying equation (7b):

$Det(I - G\otimes\Pi) = 1 - \frac{1}{2}(\pi_{mm} + \pi_{ff}) + \frac{1}{4}(\pi_{mm}\pi_{ff} + \pi_{mf}\pi_{fm}) = \frac{1}{2} + terms\,with\ z\sigma_I^{-1}$, where the terms in $z\sigma_I^{-1}$ are simplified by noting $k = \frac{1}{2}(k_m + k_f)$, $\tau = \frac{1}{2}(\tau_m + \tau_f)$, with the result that:

$$(k_m\tau_m + k_f\tau_f)/4 - (k_m\tau_m + k_f\tau_f - k_m\tau_f - k_f\tau_m)/16 = \tfrac{1}{2}[k\tau + \tfrac{1}{8}(\tau_m - \tau_f)(k_m - k_f)]$$

and $Det((\mathbf{I} - \mathbf{G}^T \otimes \Pi^T) = \frac{1}{2}(1 + \kappa z\, \sigma_I^{-1})$, where $\kappa = [k\tau + \frac{1}{8}(\tau_m - \tau_f)(k_m - k_f)]$.

$$(I - G^T\otimes\Pi^T)^{-1} = Det(I - G^T\otimes\Pi^T)^{-1}(\tfrac{3}{4} + \tfrac{1}{4}k_f\tau_f z,\ \tfrac{1}{4} - \tfrac{1}{4}k_f\tau_m z \mid \tfrac{1}{4} - \tfrac{1}{4}k_m\tau_f z,\ \tfrac{3}{4} + \tfrac{1}{4}k_m\tau_m z)$$

$$(G^T\otimes\Lambda^T) = (2\sigma_I)^{-1}(\tfrac{1}{2}i_m\tau_m,\ \tfrac{1}{2}i_f\tau_m \mid \tfrac{1}{2}i_m\tau_f,\ \tfrac{1}{2}i_f\tau_f)$$

Multiplying these matrices according to equation (7b) gives:

$$\beta = \tfrac{1}{4}i(\sigma_I + \kappa z)^{-1}(X_m^{-1}(\tau + \tau_m),\ X_f^{-1}(\tau + \tau_f))^T$$

APPENDIX 5: AN APPROXIMATION FOR Λ AND Π WHEN SELECTION IS BASED UPON BLUP

The approach is based upon the approximation to BLUP proposed by Wray and Hill (1989). For a discrete scheme was approximated by an EBV for a candidate which was a selection index constructed from six pieces of information: (i) the sire's EBV, at the time of his selection, (ii) the dam's EBV , at the time of her selection; (iii) the mean EBV of all the mates of the sire; (iv) the phenotypic mean of the paternal half-sib group, including the candidate; (v) the phenotypic mean of the full-sib family, including the candidate; and (iv) the candidate's phenotype. Denote the (co)variance matrix for these sources by $\boldsymbol{V}$, which is derived by standard index theory (Wray and Hill, 1989). The parameters used to define $\boldsymbol{V}$, and $\boldsymbol{V}$ itself, are iterated until equilibrium is reached.

The index for the sire and dam at the time of selection of the candidate offspring, and the candidate itself are given by $\boldsymbol{b}_x = \boldsymbol{V}^{-1}\boldsymbol{g}_x$ where:

for the sire: $\boldsymbol{g}_1 = (v_m,\ 0,\ 0,\ \sigma_A^2/2,\ \sigma_A^2/2,\ \sigma_A^2/2)$

for the dam: $\boldsymbol{g}_2 = (0,\ v_f,\ X_m X_f^{-1} v_f,\ X_m X_f^{-1}\sigma_A^2/2,\ \sigma_A^2/2,\ \sigma_A^2/2)$

for the candidate: $\boldsymbol{g}_3 = (v_m/2,\ v_f/2,\ X_m X_f^{-1} v_f/2,\ X_m T^{-1}\sigma_A^2/4,\ \sigma_A^2,\ X_f T^{-1},\ \sigma_A^2)$

and where σ_A^2 and σ_I^2 are the equilibrium genetic variance and index variance respectively, k_x is the variance reduction coefficient and $v_q=(1-k_q)\sigma_I^2$, T is the number of candidates of each sex. These indices form a 3 x 3 (co)variance matrix **W** where $w_{ij}=\boldsymbol{g}_i^T\boldsymbol{V}\boldsymbol{g}_j$, note $w_{33}=\sigma_I^2$.

π_{pq} assuming random mating. The terms conferring selective advantage to a parent is its estimated breeding value at the time of its selection ($\hat{A}_i$; the 'initial EBV'), its increment at the selection of its offspring ($\delta\hat{A}_i$; the 'increment'), and remaining prediction error ($\hat{e}$; the 'prediction error'). For discrete generations there are two categories, males and females. The regressions of these terms for a selected *j(p)* of sex *p* on its parent *i(q)* of sex *q* are required. It is easiest to

consider these in two parts: the regression of $\hat{A}_{j(p)}$ on $\hat{A}_i$ and $\delta\hat{A}_i$, and the regressions of $\delta\hat{A}_{j(p)}$ and $\hat{e}_{j(p)}$ on $\hat{e}_i$. Note that: (i) the prediction error of the parent is independent of the initial EBV of the progeny (otherwise the prediction of the EBV of the parent could be improved); and (ii) the increment and the prediction error of the progeny must be independent of the information on the parent at the time of progeny selection, or otherwise the EBV of the progeny could be improved.

The three terms $\hat{A}_{i(q)}$, $\delta\hat{A}_{i(q)}$, and $\hat{A}_{j(p)}$, have a (co)variance matrix before selection given by:

$$(v_q,\ 0,\ \tfrac{1}{2}v_q \mid 0,\ w_{qq} - v_q,\ w_{q3} - \tfrac{1}{2}v_q \mid \tfrac{1}{2}v_q,\ w_{q3} - \tfrac{1}{2}v_q,\ w_{33})$$

where q=1 for males and 2 for females. Appendix 2 can be followed to obtain:

$$\hat{A}_{j(p)} = \tfrac{1}{2}(1-k_p)\hat{A}_{i(q)} + (w_{q3}-\tfrac{1}{2}v_q)(w_{qq}-v_q)^{-1}(1-k_p)\delta\hat{A}_{i(q)} + \epsilon$$

Note both regression coefficients have $(1\text{-}k_p)$ in them as factors, which makes these coefficients small independent of h^2.

The regression coefficients for $\hat{e}_{i(q)}$ are less immediate. However they can be obtained from the following, where $\sigma_{a(q)}^2 = \sigma_A^2 - k_q\sigma_I^2$ is the genetic variation among parents of category q.

$$\tfrac{1}{2}\sigma_{a(q)}^2 = cov((\hat{A}_{i(q)}+\delta\hat{A}_{i(q)})+\hat{e}_{i(q)}, (\hat{A}_{j(p)}+\delta\hat{A}_{j(p)})+\hat{e}_{j(p)} \mid i(q)\ selected)$$

$$= cov(\hat{A}_{i(q)}+\delta\hat{A}_{i(q)}), \hat{A}_{j(p)} \mid i(q) selected) + cov(\hat{e}_{i(q)}\ ,\ (\delta\hat{A}_{j(p)}+\hat{e}_{j(p)}))$$

So $v_{ee} = cov(\hat{e}_{i(q)}\ ,\ (\delta\hat{A}_{j(p)}+\hat{e}_{j(p)})) = \tfrac{1}{2}\sigma_{a(q)}^2 - w_{q3}$. Selection of the offspring $j(p)$ will not effect v_{ee} since all components are independent of the initial EBV for j. Next consider the index for the offspring of $j(p)$; the covariance of the prediction error of $j(p)$ at the time of its selection, i.e. $\delta\hat{A}_{j(p)} + \hat{e}_{j(p)}$, with its unselected offspring which are included in its index is ½ of prediction error variance, and the covariance of $\hat{e}_{i(q)}$ with these unselected individuals will be $\tfrac{1}{2}v_{ee}$. Therefore define $\boldsymbol{g}_e$ = (0, 0, 0, $\tfrac{1}{2}v_{ee}$, $\tfrac{1}{2}v_{ee}$, $\tfrac{1}{2}v_{ee}$) if j is male, and (0, 0, 0, $\tfrac{1}{2}v_{ee}\mathbf{X}_m\mathbf{X}_f^{-1}$, $\tfrac{1}{2}v_{ee}$, $\tfrac{1}{2}v_{ee}$) if j is female, then cov($\hat{e}_{i(q)}$,$\delta\hat{A}_{j(p)}$) will be $\boldsymbol{g}_e^T\boldsymbol{b}_p$ and cov($\hat{e}_{i(q)}$,$\hat{e}_{j(p)}$) = v_{ee} - $\boldsymbol{g}_e^T\boldsymbol{b}_p$.

Therefore, since Var($\hat{e}_{i(q)}$) = $\sigma_{a(q)}^2 - w_{qq}$:

$$\delta\hat{A}_{j(p)} = \boldsymbol{g}_e^T\boldsymbol{b}_p(\sigma_{a(q)}^2 - w_{qq})^{-1}\ \hat{e}_{i(q)}$$

$$\hat{e}_{j(p)} = (½\sigma_a^2 - w_{q3} - \boldsymbol{g}_e^T \boldsymbol{b}_p)(\sigma_{a(q)}^2 - w_{qq})^{-1} \hat{e}_{i(q)}$$

λ_{pq} *assuming random mating.* Following Appendix 2: the regressions are: for $\hat{A}_{i(q)}$, $½ i_p \sigma_i^{-1}$; for $\delta\hat{A}_{i(q)}$, $½ i_p \sigma_I^{-1} (w_{q3} - ½(1-k_q)w_{33})/(w_{qq} - (1-k_q)w_{33})$; and 0 for $\hat{e}_{i(q)}$.

Covariances between selective advantages and Mendelian terms. Define $\boldsymbol{g}_a$ = (0, 0, 0, $½h_0^2 X_m/T$, $½h_0^2 X_m/X_f$, $½h_0^2$), then the mean of the Mendelian sampling terms for selected ancestors is $E[a_{i(q)}] = i_q \boldsymbol{g}_a^T \boldsymbol{b}_3 / \sigma_I^2$ and . Therefore $cov(a_{i(q)}, \delta\hat{A}_{i(q)} + \hat{e}_{i(q)}) = v_{aa} = ½h_0^2$ - $cov(a_{i(q)}, \hat{A}_{i(q)})$. To determine the covariances for $\delta\hat{A}_{i(q)}$ and $\hat{e}_{i(q)}$, consider the covariances of the 6 information sources defined by Wray and Hill, with v_{aa}: for q male; and for q female, $\boldsymbol{g}_a^*$ = (0, 0, 0, $½v_{aa} X_m/X_f$, $½v_{aa}$, $½v_{aa}$). Then $\mathrm{cov}(a_{i(q)}, \delta\hat{A}_{i(q)}) = \mathbf{g}_a^T \boldsymbol{b}_q$, and $\mathrm{cov}(a_{i(q)}, \hat{e}_{i(q)}) = v_{aa} - \mathbf{g}_a^T \boldsymbol{b}_q$,

Example (X_m=20, X_f=40, h^2=0.4, 8 offspring per litter). The matrices are presented in the following row and column order: for dimension 6 the order is ($\hat{A}_{i(m)}$, $\delta\hat{A}_{i(m)}$, $\hat{e}_{i(m)}$, $\hat{A}_{i(f)}$, $\delta\hat{A}_{i(f)}$, $\hat{e}_{i(f)}$), whilst for dimension 2 the order is males then females.

$\mathbf{G}\otimes\Lambda$ = (1.16, 2.07, 0, 1.16, 2.71, 0 | 0.90, 1.60, 0, 0.90, 2.09, 0)

$\mathbf{G}\otimes\Pi$ = (0.046, 0.082, 0, 0.046, 0.107, 0 | 0, 0, 0.049, 0, 0, 0.053 | 0, 0, 0.081, 0, 0, 0.087
| 0.060, 0.108, 0, 0.060, 0.141, 0 | 0, 0, 0.032, 0, 0, 0.035 | 0, 0, 0.098, 0, 0, 0.106)

$cov(a_{i(q)}, A_{i(q)})$ = (0.011, 0.014), $cov(a_{i(q)}, \delta A_{i(q)})$ = (0.054, 0.035), $cov(a_{i(q)}, \hat{e}_{i(q)})$ = (0.088,0.107)

Table 1. Approximations to rates of gain for general half-sib indices using formulae developed in the text, either with initial or equilibrium parameters. The schemes assumed 20 male and 40 female parents respectively with eight offspring per litter. Initial heritability was assumed equal to 0.4. (z is as defined in the text, the product of the accuracy and the genetic standard deviation).

Index			Estimates of rates of gain using initial parameters			Equilibrium
b_1	b_2	b_3	iz[1]	Equation (13)[2]	Equation (14)[3]	Parameters
1	0.5	0	0.353	0.344	0.344	0.344
1	1	1	0.584	0.444	0.466	0.480
1	1.5	2	0.650	0.437	0.481	0.511
1	3	6	0.658	0.403	0.457	0.502

(1) Uses z obtained in the base generation.

(2) Uses both z and σ_I obtained in the base generation.

(3) Uses σ_I obtained in the base generation.

Table 2. The partition of genetic gain into unique and independent contributions arising from within and between female families, and within and between male families. The schemes assumed 20 male, and 40 female parents respectively with eight offspring per litter. Initial heritability was assumed equal to 0.4.

Index			Total	Males		Females	
b_1	b_2	b_3	Gain	Within	Between	Within	Between
1	0.5	0	0.344	0.176	0.008	0.136	0.025
1	1	1	0.480	0.172	0.086	0.133	0.089
1	1.5	2	0.511	0.153	0.122	0.118	0.117
1	3	6	0.502	0.096	0.176	0.074	0.155

Table 3. A comparison of simulated and predicted responses for selection using BLUP. The scheme breed 20 male and 40 female parents with six offspring per litter, with h^2=0.4. The simulations were for 400 replicates and the s.e. derived from replication error are given in parentheses.

	Simulation		Equilibrium		Base parameters	
	Male	Female	Male	Female	Male	Female
$\beta(\hat{A})$	0.056 (0.004)	0.032 (0.002)	0.057	0.029	0.057	0.028
$\beta(\delta\hat{A})$	0.103 (0.002)	0.068 (0.001)	0.103	0.067	0.102	0.067
$\beta(\hat{e})$	0.011 (0.002)	0.006 (0.001)	0.012	0.006	0.011	0.006
ΔG (within)	-	-	0.135	0.104	0.134	0.103
$\Delta G(\hat{A})$	-	-	0.011	0.016	0.011	0.016
$\Delta G(\delta\hat{A})$	-	-	0.105	0.092	0.105	0.094
$\Delta G(\hat{e})$	-	-	0.019	0.026	0.018	0.024
Total	0.507 (0.002)		0.508		0.498	

Table 4. A comparison of simulated and predicted responses for selection with maternally imprinted variation. The scheme breed 20 male and 40 female parents with six offspring per litter, with h^2=0.4. The simulations were for 400 replicates and the s.e. derived from replication error are given in parentheses.

	Sex	Simulation	Prediction using parameters from:	
			Equilibrium	Base
β (x100) *	male	0.69 (0.041)	0.73	0.71
	female	1.07 (0.016)	1.09	1.06
ΔG (between	male	0.023 (0.013)	0.024	0.023
families)	female	0.073 (0.014)	0.073	0.070
ΔG (within	male	0.076 (0.0012)	0.077	0.075
families)	female	0.055 (0.0008)	0.056	0.055
Total ΔG		0.229 (0.0014)	0.230	0.223

*: The predictions of β for A^- and A^+ were identical, and simulations were not significantly different; therefore results have been pooled.

Table 5. The time course of expected contributions from an individual female parent of age one at t=0. The breeding scheme has mass selection with 20 male parents of age 1, 40 female parents at ages one and three (20 of each age), eight offspring per litter, and heritability 0.4. The expected contribution is $c(t) + b(t)(A_i - \bar{A})$.

	To males age one		To females age one		To females age three	
Time	c(t)	b(t)	c(t)	b(t)	c(t)	b(t)
t=1	0.0167	0.0140	0.0167	0.0140	0	0
t=2	0.0151	0.0157	0.0151	0.0157	0	0
t=3	0.0132	0.0146	0.0132	0.0146	0.0167	0.0140
t=6	0.0148	0.0168	0.0148	0.0168	0.0145	0.0160
t=10	0.0149	0.0170	0.0149	0.0170	0.0149	0.0170

Paper 33

DYNAMIC SELECTION PROCEDURES FOR CONSTRAINED INBREEDING AND THEIR CONSEQUENCES FOR PEDIGREE DEVELOPMENT

B. Grundy[1], B. Villanueva[1] and J. A. Woolliams[2]

[1]Scottish Agricultural College, West Mains Road, Edinburgh EH9 3JG, United Kingdom

[2]Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, United Kingdom

Short Title: Selection procedures for constrained inbreeding

Proofs to be sent to: B. Grundy, Scottish Agricultural College, Genetics and Reproduction Department, Animal Biology Division, Bush Estate, Penicuik, EH26 0PH, Scotland UK. Tel: 0131 535 3225; Fax: 0131 535 3121; e-mail: b.grundy@ed.sac.ac.uk

SUMMARY

A novel selection algorithm for maximising genetic response while constraining the rate of inbreeding is presented. It is shown that the proposed method controls the rate of inbreeding by maintaining the sum of squared genetic contributions at a constant value and represents an improvement on previous procedures. In order to maintain a constant rate of inbreeding the contributions from all generations are weighted equally and this is facilitated by modifying the numerator relationship matrix. By considering the optimisation of the contributions of many generations the initial mating proportions (the genetic contributions to the next generation) are not equal to their long-term values, but are set equal to the expected long-term contributions given the current information. This is confirmed by the regression of the long-term contributions on the assigned mating proportions being close to one. The gain obtained from the selection algorithm is compared to the maximum theoretical genetic gain under constrained inbreeding. It is concluded that this theoretical upper bound is in general unattainable, but from this a concept of genetic efficiency in terms of resources and constraints is derived.

1. INTRODUCTION

The problem of optimising genetic progress with constrained rates of inbreeding has only recently been explored, following the development of a coherent underpinning theory for predicting rates of inbreeding under selection (Wray & Thompson, 1990). Optimal designs for maximising gain with constrained inbreeding have been developed with mass selection (Villanueva, Woolliams & Gjerde, 1996), and with sib-indices (Villanueva & Woolliams, 1997). However, there are limitations: (i) these designs are special cases, and do not, for example, cover selection using BLUP; and (ii) they are tools for designing schemes *a priori*, but do not offer guidance for selection decisions in practice which involve a particular given set of candidates with performance and pedigree records.

Unlike the advances in the deterministic predictions for genetic gain and inbreeding, the development of dynamic selection algorithms for designing schemes *a posteriori* (Woolliams & Meuwissen, 1993; Wray & Goddard, 1994; Brisbane & Gibson, 1995) has proceeded largely without the application of genetic contribution theory (Woolliams & Thompson, 1994). The approach employed in the design of these algorithms is to describe the problem of a constrained maximum gain as a quadratic programming problem. The gain is maximised step by step, optimising progress one generation in the future. Where the constraint has been on the rate of inbreeding the selection decision has involved consideration of the relationships between successful candidates, and has allowed different mating proportions for selected individuals. Meuwissen (1997) developed an algorithm that showed how to obtain explicit solutions for this problem, which are near-optimal using the numerator relationship

matrix (A). The shortfalls of these procedures are that (i) they do not achieve a constant rate of inbreeding over several generations of selection; and (ii) because of a lack of accompanying theory it is unclear if consideration of only the next generation provides the optimal, feasible solution to the long-term problem.

An approach to understanding the theoretical problem is provided by the concept of long-term genetic contributions (James & McBride, 1958). Woolliams & Thompson (1994) developed this approach to show that the long-term rates of gain and inbreeding that were attributable to a generation could be described in a unified way by contributions. Furthermore they showed that the recurrent selection problem constrained by rates of inbreeding could be recast in a form that was analogous to a previously solved problem of optimal allocation of clones in one generation (Bondesson, 1989).

This paper has four objectives. Firstly, to improve existing methods for maximising gain while restricting the rate of inbreeding by introducing modifications to existing algorithms (using an augmented numerator relationship matrix ($\mathbf{A}^*$) and an associated constraint). Secondly, to examine the relationship of the optimal mating proportions in dynamic selection procedures with long-term contributions in order to describe the process by which individual contributions influence the way the pedigree develops. Thirdly, to extend the results of Woolliams & Thompson (1994) for deriving explicitly a theoretical upper bound to genetic progress as a function of the constraints imposed and the resources available. Finally, to develop a concept of genetic efficiency by comparing the results of applying the modified dynamic algorithm with the theoretical

upper bound.

2. METHODS

(i) Simulation models

Stochastic computer simulations were used to model the effects of dynamic selection routines on the way the pedigree develops and to compare rates of gain and inbreeding from using different selection algorithms. Populations with discrete generations were evaluated over 20 generations. An additive infinitesimal model (Bulmer, 1971) was considered. The trait under selection was assumed to be of known heritability (h^2) and the genetic evaluation was carried out using an animal model BLUP. True breeding values of unrelated individuals in the base population ($t = 0$) were obtained from a normal distribution with mean zero and variance h^2. Phenotypic values were obtained by adding a normally distributed environmental component with mean zero and variance $(1 - h^2)$. In each subsequent generation, the optimum number of sires and dams were selected as described below. The number of offspring born per generation was 100 (50 males and 50 females). True breeding values of the offspring were generated as half the sum of the true breeding values of the animal's sire and dam plus a Mendelian sampling term taken from a normal distribution with mean zero and variance $(½)[1 - ½(F_s + F_d)]h^2$, where F_s and F_d are the inbreeding coefficients of the sire and the dam, respectively. The rates of inbreeding and response were calculated for each generation.

Three different dynamic selection algorithms were employed, all based upon the algorithm presented by Meuwissen (1997). For each algorithm, the number of sires

and dams and their contributions were optimised each generation of selection in order to maximise the rate of gain subject to a constraint on the rate of inbreeding.

(a) *Algorithm I*

This was as described by Meuwissen (1997). In brief, the algorithm uses the estimated breeding values and the numerator relationship matrix ($\mathbb{A}$) to identify an optimum mating proportion for each individual i at generation t ($c_{i,t}$), where $c_{i,t} = 0$ implies that the individual is not required for breeding. It achieves this by maximising

$$f(\mathbf{c}) = \mathbf{c}_t^{\mathrm{T}}\mathbf{g}_t - \lambda\mathbf{c}_t^{\mathrm{T}}\mathbf{A}_t\mathbf{c}_t$$

where $\mathbf{c}_t$ is the vector of mating proportions of selection candidates at generation t (i.e. genetic contributions to the next generation), $\mathbf{g}_t$ is the vector of estimated breeding values, $\mathbf{A}_t$ is the numerator relationship matrix of candidates, and λ is a Lagrangian multiplier. Constraints are imposed so that mating proportions are positive (i.e. $c_{i,t} \geq 0$), and sum to 1 (i.e. $\sum_i c_{i,t} = 1$). The Lagrangian multiplier is chosen so as to achieve the constraint $\mathbf{c}_t^{\mathrm{T}}\mathbf{A}_t\mathbf{c}_t \leq 2t\Delta F$, where ΔF is the intended rate of inbreeding and t is the generation number.

For an individual with mating proportion $c_{i,t}$ the desired number of offspring was $200c_{i,t}$ (since 100 was the total number of candidates per generation and the contributions sum to a half for each sex). The actual number was achieved through a process of: (i) reducing the desired real number to the largest integer below it and then allocating a mate at random to produce a single offspring per mating; and (ii) adding further offspring to parents with the greatest deviations from the desired number until the

number of offspring sum to 100, again allocating mates at random. The sex of the offspring was randomly determined.

(b) *Algorithm II*

Algorithm I controls the increase in average coancestry and hence it constrains the absolute increase in inbreeding coefficient from time $t-1$ (F_{t-1}) to time t (F_t), rather than the rate of inbreeding. This increase in the inbreeding coefficient approximates ΔF only when F_{t-1} is small. In Algorithm II a straightforward modification of Algorithm I to achieve the restriction on ΔF was applied by setting the constraint, $C_t = C_{t-1} + 2\Delta F(1-F_{t-1})$, where F_{t-1} was calculated *a priori* as $\frac{1}{2}C_{t-1}$ and $C_0 = 0$.

(c) *Algorithm III*

In contrast to the previous algorithms, Algorithm III controls the total increase in squared genetic contributions each generation to constrain ΔF. This was achieved by modifiying Algorithm I by replacing **A** with an augmented matrix **A***. The augmented matrix for individuals in the generation $t+1$ was calculated as

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

where **D** is diagonal with elements equal to ½ and $\mathbf{Z}_t$ is a gene flow matrix (Hill, 1974) identifying parents of generation $t+1$. With the offspring in rows and parents in columns, the elements of $\mathbf{Z}_t$ are either ½ (parents) or 0 (otherwise) (Thompson, 1977; Wray & Thompson, 1990). For $t = 0$ (base population), $\mathbf{A}^*_0 = \mathbf{I}$. By augmenting **A** the diagonal terms are no longer scaled by one minus the average inbreeding of the parents and hence the base is in effect re-defined each generation. Therefore each

generation can used to constrain ΔF to a constant value as the differential treatment of generations arising from this dependence on the average inbreeding of the parents is removed. A proof that $\mathbf{c}_t^{\mathrm{T}}\mathbf{A}_t^*\mathbf{c}_t$ can be used to constrain the rate of inbreeding is presented in Appendix A.

The constraint applied in this algorithm is $\mathbb{c}_t^{\mathrm{T}}\mathbb{A}_t^*\mathbb{c}_t \leq t\Delta C$, where ΔC is set to $2\Delta F\,[1-3\Delta F+12(\Delta F)^2]$; this is approximately equal to ΔF for small values of ΔF but is marginally greater than ΔF for larger ΔF, due to second order effects (see Appendix B). Henceforth $\mathbf{A}_t^*$, $\mathbf{A}_t$, and $\mathbf{c}_t$ will be abbreviated as $\mathbf{A}^*$, $\mathbf{A}$, and $\mathbf{c}$, respectively.

(d) *Relationship between mating proportions and long-term contributions*

The total number of parents selected each generation was obtained along with the 'effective' number (N) of each sex which was derived from $(4\mathbf{c}^{\mathrm{T}}\mathbf{c})^{-1}$ (Robertson, 1965). For each replicate using $\mathbf{A}^*$, the long-term contributions ($\mathbf{r}$) were calculated using the program used by Woolliams & Mäntysaari (1995) for individuals born at generation 15. The long-term contribution of an individual i was defined as the proportion of genes in generation 20 that derived from that ancestor. The regression of r_i on c_i was calculated for each replicate using only those individuals for which $c_i > 0$. In addition, $\mathbb{r}^{\mathrm{T}}\mathbb{r}$, $\mathbb{c}^{\mathrm{T}}\mathbb{c}$ and their ratio were calculated for ancestors in generation 15. Generation 15 was chosen since this is the last generation whose contributions are near convergence by the end of the simulation ($t = 20$).

To obtain the regression of $\mathbb{r}$ on $\mathbb{c}$, the results for generation 15 were subdivided by sex, replicate and the number of parents selected. The regression coefficients and their

associated standard errors were calculated within each category and were pooled using weights inversely proportional to the sampling variance. An analysis of variance found no significant differences in slope due to number of parents.

(e) *Parameters*

The schemes were run for a range of heritabilities (0.01, 0.25 and 0.99) and constraints on the rate of inbreeding (0.00625, 0.025 and 0.05). A minimum of 100 replicates were run for each combination and results presented are averages over replicates.

(ii) Upper bounds for rate of progress and efficiency

Bondesson (1989) solved an allocation problem concerning the planting of optimal proportions (k_i) of clones with known genetic values (A_i) to maximise gain ($\sum k_i A_i$) in a single crop with a constraint on genetic diversity (namely $\sum k_i^2 \leq \gamma^{-1}$). Here γ has a lower bound of one attained when only one clone is planted throughout the crop. This problem is equivalent to a recurrent selection problem with constrained inbreeding if we consider Mendelian sampling terms (a_i) rather than genetic values and long-term contributions (r_i) rather than clonal proportions. For infinite populations (Bondesson, 1989) the solution to the problem of obtaining maximum genetic gain under constrained diversity was obtained by maximising

$$\int a\, r(a)\varphi(a)\mathrm{d}a$$

where the intergration is bounded by $-\infty$ and $+\infty$, subject to the constraints (1):

(i) $r(a) \geq 0$ (i.e. all contributions are non negative);

(ii) $\int r(a)\varphi(a)\, \mathrm{d}a = 1$ (i.e. the total contribution over all individuals is one);

(iii) $\int [r(a)]^2\varphi(a)\, \mathrm{d}a = \gamma^{-1}$ (i.e. the diversity is constrained)

where $\varphi(a)$ is the density function for the Mendelian sampling terms and $r(a)$ is the total long-term contribution in the population for individuals with Mendelian sampling term a.

When finite population sizes and constraints on rates of inbreeding are considered then the number of individuals in an interval a to $a+\delta a$ is $2T\varphi(a)\delta a$ (where T is total number of candidates per sex, and the 2 arises from considering both sexes). The problem can be recast to provide the optimum solution to the recurrent selection problem considered here of maximising ΔG with a constraint on ΔF as $\Delta G = 2\ T\ E\ [r_i\ a_i]$ and $\Delta F = 2\ T\ E\ [r_i^2]$ (Woolliams & Thompson, 1994). Here we will assume $\Delta F = \frac{1}{2}\Delta C$, as for the majority of cases the difference between the two is small. Thus the solution to the constrained maximisation problem is obtained by maximising

$$\Delta G = \int 2T\, a\, r(a)\varphi(a)\mathrm{d}a$$

subject to the constraints

$$\text{(i) } r(a) \geq 0;\ \text{(ii) } \int 2T\, r(a)\varphi(a)\ \mathrm{d}a = 1;\ \text{(iii) } \int 2T\, [r(a)]^2\varphi(a)\ \mathrm{d}a = 4\Delta F$$

Multiplying through the third constraint by $2T$ and substituting $r^*(a) = 2Tr(a)$ we recover the form of equations (1) with $\gamma^{-1} = 8T\Delta F$. Since the minimum rate of inbreeding with T candidates per sex is $(8T)^{-1}$, the constraint that $\gamma \leq 1$ is satisfied. Therefore the maximum ΔG is identical to the solution of Bondesson (1989) with γ replaced by $(8T\Delta F)^{-1}$.

Assuming a normally distributed Mendelian sampling term, and expressing gain in terms of i (mean deviation of individuals with values exceeding the truncation point) and x (deviation of the truncation point from the mean), the expression of Bondesson

(1989) for the maximum theoretical (ideal) gain (ΔG_{ideal}) can be rewritten for recurrent selection as

$$\Delta G_{ideal} = (i-x)^{-1}$$

and the values for i and x are such that γ is the solution of

$$\gamma = 2p(i-x)^2(1+x^2-ix)^{-1}$$

where p is the proportion selected. The solution for x can be found by using the Newton-Raphson method.

The above expression for ΔG_{ideal} assumes a standard deviation of Mendelian sampling terms of one. Making this expression more general the maximum theoretical genetic gain per unit time can be expressed in terms of base phenotypic standard deviation, resources available (i.e. number of candidates) and risk (i.e. rate of inbreeding) as

$$\Delta G_{ideal} = i\,(k)^{-1}\sqrt{\tfrac{1}{2}h^2}$$

where $k = i(i-x)$. This equation can be used to check if the proposed method (Algorithm III) not only constrains ΔF but also maximises ΔG. The genetic response (ΔG_{obs}), averaged over replicates, was estimated from the sum of the products of the Mendelian terms and the long term contributions of the candidates in generation 3 (i.e. $\Delta G_{obs} = \sum r_{i,3}a_{i,3}$). Generation 3 was chosen for this evaluation as the impact of the Bulmer effect has largely taken place and yet the reduction in genetic variance associated with the mean level of inbreeding is negligible. The ratio of ΔG_{obs} to ΔG_{ideal} can be considered as a measure of relative genetic efficiency.

(iii) Relative efficiencies of other selection procedures

Under the same constraints on resources and the rate of inbreeding, the rates of gain obtained using Algorithm III were compared to deterministic predictions of maximum gain with optimised mass selection and with optimised sib-indices. These deterministic predictions used the methods of Villanueva & Woolliams (1997), but with one exception in that the Mendelian sampling variance was not reduced due to inbreeding. The models for mass and sib-index selection assumed equal full- and half-sib family sizes and constant numbers selected per generation with hierarchical mating where appropriate (i.e. when the optimum mating ratio was greater than one).

3. RESULTS

(i) Augmented relationship matrix

A comparison of rates of inbreeding and response obtained with Algorithm I (using $\mathbb{A}$), Algorithm II (using $\mathbf{A}$ and a modified constraint) and Algorithm III (using $\mathbf{A}^*$ and a modified constraint) is shown in Table 1 for a heritability of 0.25 and a desired rate of inbreeding of 0.025. The sum of the squared contributions ($\mathbf{r}^T\mathbf{r}$) and consequently the rate of inbreeding ($¼\mathbf{r}^T\mathbf{r} = \Delta C$) were maintained at their predefined levels throughout the period of selection with $\mathbf{A}^*$ and the modified constraint but not with the standard $\mathbb{A}$. The squared contributions increased over time with Algorithm I, and decreased with Algorithm II. The optimal numbers selected were the same for both sexes and were also constant with $\mathbb{A}^*$ but they declined (Algorithm I) or increased (Algorithm II) over time when $\mathbb{A}$ was used. Despite the inability to constrain the rate of inbreeding when using $\mathbb{A}$, there was little difference in the rate of response achieved by generation 20 amongst the three methods. Algorithm II gave slightly lower responses and a lower rate of inbreeding than the other algorithms. However it would be expected that if

Algorithm III was constrained to the same lower rate of inbreeding the response achieved would be greater than that obtained with Algorithm II. When in Algorithm II the F_{t-1} term in the constraint was calculated from the data (rather than calculated *a priori*) the results were very similar to those shown in table 1 (results not shown).

Different schemes which covered a broad range of heritabilities (0.01, 0.25 and 0.99) and possible constraints (given the number of candidates available for selection) on ΔF (0.05, 0.025 and 0.00625) were simulated with Algorithm III. In all these cases the desired ΔF was achieved.

(ii) Relationship of **r** with **c**

The regression of the long-term contributions from animals born in generation 15 to generation 20 on the original mating proportions assigned to animals born in generation 15 is shown in Table 2 for a range of heritabilities and constraints on the rate of inbreeding. The regression of r_i on c_i was close to one for all the cases considered. As heritability decreases so the regression coefficients became smaller.

(iii) Relationship of $\mathbf{r}^T\mathbf{r}$ with $\mathbf{c}^T\mathbf{c}$

The ratio of $\mathbf{r}^T\mathbf{r}$ to $\mathbf{c}^T\mathbf{c}$ is presented in Table 3. The ratio decreased with increased heritability and increased severity of the restriction on the rate of inbreeding. As heritability tended to one, and as the restriction on the rate of inbreeding became more severe, so the ratio tended to one.

(iv) Efficiency of genetic gain

Table 4 shows the efficiency of the genetic response obtained with Algorithm III (ΔG_{obs}) relative to the deterministic predictions of the ideal genetic response (ΔG_{ideal}). When the heritability is close to unity (h^2 = 0.99) the responses obtained with Algorithm III as a proportion of the ideal ranged between 81% and 87%. As the heritability decreased, there was a dramatic drop in efficiency, approximately related to $\sqrt{h^2}$, with little variation across the different constraints on the rate of inbreeding.

There was variation in efficiency over replicates, and in some replicates the gain exceeded the upper bound (ΔG_{ideal}) (i.e. the efficiency was greater than one). For example when the heritability is close to one (h^2 = 0.99) the gain derived from the ideal solution was exceeded in 3%, 12% and 16% of the replicates, for rates of inbreeding of 0.00625, 0.025 and 0.05, respectively. Therefore it is possible for a particular generation within a replicate to produce more gain than the upper bound, but this is due to the random sampling of matings (even though the average relationship $\mathbf{c}_t^T\mathbf{A}_t^*\mathbf{c}_t$ is $\leq C_t$), an effect that is more prominent when the number of mates is small (e.g. ΔF = 0.05). However, a persistent breaking of the upper bound can not be sustained over several generations as these matings also yield more inbreeding.

(v) Comparison of Algorithm III with other selection procedures

Deterministic predictions for the maximum predicted gains with constrained inbreeding with mass (Villanueva, Woolliams & Gjerde, 1996), and sib-indices (Villanueva & Woolliams, 1997) selection were calculated for a range of constraints on the rate of inbreeding and heritabilities. The models were slightly modified to produce an asymptotic rate of response (i.e. Bulmer equilibrium) by ignoring the reduction in

Mendelian sampling variance by inbreeding. The asymptotic responses obtained from Algorithm III (which uses BLUP), mass and sib-index selection are presented in Table 5. The efficiencies of constrained mass selection with respect to Algorithm III were always lower than one, ranging from 0.65 to 0.91. Sib-indices were more efficient than mass selection, with efficiencies with respect to Algorithm III ranging from 0.71 to 1.

4. DISCUSSION

It has been shown that through the use of a modified relationship matrix ($\mathbf{A}^*$) the rate of inbreeding can be restricted at a predefined level for successive generations of selection. This constant rate is achieved by setting the mating proportions ($\mathbf{c}$) to values consistent with the expectations of their long-term contributions ($\mathbf{r}$) at the time of selection, which in turn is indicated by the regression of long-term contributions on mating proportions being close to one. However, even when the heritability is close to one, $\mathbf{c}^T\mathbf{c}$ does not equal $\mathbf{r}^T\mathbf{r}$. This can be interpreted as an inability to achieve the ideal solution where $\mathbf{c}$ equals $\mathbf{r}$ and $\mathbf{c}^T\mathbf{c}$ equals $\mathbf{r}^T\mathbf{r}$. The impact of this failure to achieve the ideal solution was shown to be primarily dependent on heritability and largely independent of the constraint on inbreeding. A theoretical response associated with the ideal solution was derived and hence the loss of response was quantified as an efficiency, defined as the ratio of the observed response to the ideal theoretical response.

(i) <u>The use of A* to control pedigree development</u>

Rates of inbreeding are determined by the contributions of the current generation and the contributions of ancestral generations that have yet to converge. In this respect all

generations have equal importance in controlling rates of inbreeding and contributions from all these generations need to be given equal weight. With Algorithm I, the weighting factor associated with each generation in **A** is unequal since the contributions of later generations are down-weighted by a term that is approximately $\frac{1}{2}(1-F_{t-1})$ where F_{t-1} is the average inbreeding coefficient of the preceding generation. As a result the sum of the squared contributions ($\mathbf{r}^{T}\mathbf{r}$) of later generations are allowed to inflate, thereby increasing ΔF (as observed in Meuwissen, 1997). Removal of the terms containing F_{t-1} in the augmented $\mathbf{A}^{*}$ is a natural approach to the problem arising from consideration of genetic contributions and avoids the problems of Meuwissen (1997). The weighting factor for the contributions at any generation remains stable over time with the value ½.

An alternative approach using the standard **A** was examined using Algorithm II and found to be unsatisfactory. The constraint $C_t = C_{t-1} + 2\Delta F$ of Algorithm I was replaced by $C_t = C_{t-1} + 2(1-F_{t-1})\Delta F$. This continuous modification to the increment applied to the constraint over time ($1-F_{t-1}$) ensures the value of the constraint is consistent with the predicted course of ΔF over time. As the constraint was met ($\mathbf{c}_t^{T}\mathbf{A}_t\mathbf{c}_t \leq C_t$) the failure to constrain ΔF can be attributed to the way in which the simultaneous optimisation of contributions of both the current and previous generations occurs when using $\mathbb{A}$. In this case the weighting factors for individual contributions is $\frac{1}{2}\left(1-\overline{F}\right)$, where $\overline{F}$ is the average inbreeding of the parents of the individual ancestor (i.e. it is not precisely the average inbreeding for the parental generation; F_{t-1}). Hence, the problem involves optimising $\mathbf{r}^{T}\mathbb{D}\mathbf{r}$, where $\mathbb{D}$ is a diagonal matrix of the terms

$\frac{1}{2}\left(1-\overline{F}\right)$, and the constraint ($C_t$) can be met by varying both components **r** and **D**, with the consequence that the observed ΔF which is proportional to $\mathbf{r}^T\mathbf{r}$ is not as intended.

(ii) Genetic Efficiency

Although Algorithm III is able to constrain the rate of inbreeding to a predefined level, the question arises as to whether maximisation of the function $f(\mathbf{c}) = \mathbf{c}_t^T\mathbf{g}_t - \lambda\mathbf{c}_t^T\mathbf{A}_t{}^*\mathbf{c}_t$ yields the maximum possible genetic gain. The ideal response under constrained inbreeding (ΔG_{ideal}) was derived assuming that i) the genes from the current generation can be completely mixed throughout all individuals in the next generation and ii) that breeding values are known without error. Even when the accuracy is one, with a dioecious population and within a single generation this mixing process can not be achieved, as parental genes are present only in their offspring. Hence, the selection of the current generation is not independent of previous generations with contributions still converging. Furthermore, this dependence is increased when the accuracy is less than one. However, the selection decisions obtained from Algorithm III are not achieved by considering only the current generation but by the simultaneous optimisation of the contributions from the current and previous generations (Appendix C).

Therefore Algorithm III may be viewed as an empirical upper bound to response (ΔG_{ub}), in contrast to the over-prediction of the ideal response of ΔG_{ideal}. The deviation of ΔG_{ideal} from ΔG_{ub} is dependent on the population structure, and for the case where both sexes are measured once before selection (Table 4 and results not

shown for $h^2 = 0.5$), $\Delta G_{ub} \approx \rho\ \Delta G_{max}$ where ρ is the accuracy of the Mendelian sampling term for a selected individual. We have described in Appendix D an improved prediction that was within 0.01 for all schemes in Table 4.

As an aside, since $\mathbf{c} = \mathbf{r}$ represents an ideal outcome, it is possible to view $(\mathbf{c}^T\mathbf{c})(\mathbf{r}^T\mathbf{r})^{-1}$ as a measure of the efficiency of the scheme or the efficiency of the dispersal of genes throughout the population. The results of this study show that contrary to the speculation of Woolliams & Thompson (1994), this ideal appears unattainable even with a heritability of one, except in the special case of the extreme lower bound for inbreeding (0.0025 in this example). In this special case there is no selection and each parent is required to be replaced by two offspring, (there is no variation in family size) and hence $\mathbf{c}$ is identical to $\mathbf{r}$ (results not shown). The results shown by Meuwissen (1997) tend to obscure the lack of correspondence between $\mathbf{c}^T\mathbf{c}$ and $\mathbf{r}^T\mathbf{r}$ since the apparent close relationship between $\mathbf{c}^T\mathbf{c}$ and $\mathbf{r}^T\mathbf{r}$ in his results arises from the bias produced by using $\mathbf{A}$ in which contributions from later generations are inflated and, consequently, failing to constrain ΔF to be constant over multiple generations.

(iii) The use of selected individuals

With Algorithm III the degree of relationship between the mating proportions ($\mathbf{c}$) and breeding values ($\mathbf{g}$), termed the usage solution, is dependant on the constraint imposed on ΔF. The form of the distribution of mating proportions is summarised by the linear regression of $\mathbf{c}$ on $\mathbf{g}$, which is a perfect linear regression for the ideal genetic gain (Bondesson, 1989). The general form of the solution with Algorithm III with regard to the constraint on ΔF is that as a less severe constraint is imposed there is: i) an

increase in the value of the intercept as fewer individuals are used; ii) an increase in the slope of the line as usage of selected individuals becomes more unequal and; iii) an increase in the goodness of fit of the regression line as more emphasis is placed on breeding value.

The general form of this distribution of mating proportions differs from those solutions obtained with truncation selection where all individuals with breeding values (g) above the truncation point are used uniformly. Comparison of these two forms of solution indicates that truncation selection is less efficient (i.e. yields less response) than Algorithm III when schemes are compared at the same rate of inbreeding (Toro & Nieto, 1984).

(iv) Comparison with other studies

In recent years several selection methods have been developed to control inbreeding by placing a direct constraint on either the cumulative inbreeding (Wray & Goddard, 1994; Brisbane & Gibson, 1995) or the rate of inbreeding (Meuwissen, 1997; Villanueva & Woolliams, 1997). These procedures are more efficient than standard truncation selection. The procedure of Villanueva & Woolliams (1997) was for sib-indices which yielded empirical efficiencies that were lower than those obtained with Algorithm III, as might be expected given the lower accuracy of sib-index selection. However, in addition, their scheme is static, with constant numbers selected and equal full-sib family sizes which reduces the possible response as all selected individuals are assigned the same mating proportions.

Dynamic procedures have also been proposed (Wray & Goddard, 1994; Brisbane & Gibson, 1995; Meuwissen, 1997) which are based on optimising the number and usage of parents each generation, an idea first proposed by Toro & Nieto (1984). The procedure of Wray & Goddard (1994) was aimed to maximise the long term selection response, which was obtained by weighting the rate of inbreeding against the selection differential. Their algorithm was sub-optimal although the contribution of the parents given their selection was determined optimally. Wray & Goddard (1994) demonstrate an advantage of their procedure over standard BLUP of 7% by generation 30, with h^2 = 0.4, and selection only on males. Meuwissen (1997) reported a greater advantage in response over BLUP (up to 60%) at the same rate of inbreeding. The results presented here (Table 1) indicate that similar increases in response will be achieved with Algorithm III.

For the practical application of the method in animal breeding, a model that includes reproductive limits and overlapping generations needs to be considered. An extension to include reproductive limits can be accommodated within the optimisation by the inclusion of additional constraints and achieved either by considering fixed contributions (Meuwissen, 1997) or through the use of an Evolutionary Algorithm (Grundy, Kinghorn, Villanueva, & Woolliams, 1997). Additional constraints can also be used within the procedures to account for overlapping generation structure (Grundy, Villanueva & Woolliams, 1997; Meuwissen, 1998) and as shown by Meuwissen (1998) yield extra benefit compared to truncation selection.

ACKNOWLEDGEMENTS

This work was funded by the BBSRC and the EU. SAC also receives funds from the Scottish Office, Agriculture, Environment and Fisheries Department. Work at the Roslin Institute receives funds from MAFF.

REFERENCES

Bondesson, L. (1989). Mathematical principles of optimal utilisation of clones under diversity constraints. *Swedish University of Agricultural Sciences, Section of Forest Biometry, Report 27.*

Brisbane, J.R. & Gibson, J.P. (1995). Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theoretical and Applied Genetics* 91, 421-431.

Bulmer, M.G. (1971). The effect of selection on genetic variability. *The American Naturalist* **105**, 201-211.

Grundy, B., Villanueva B. & Woolliams, J.A. (1997). Maximising response with a predefined rate of inbreeding for overlapping generation structure. *Proceedings of British Society of Animal Science.* Scarborough, pp 27.

Grundy, B., Kinghorn, B., Villanueva, B. Woolliams, J. A. (1997). A breeding programme for in-situ genetic conservation of livestock. *British Society of Animal Science Conference and Workshop on 'The Potential Role of Rare Livestock Breeds in UK Farming Systems'.* Appleby, pp 41.

Hill, W.G. (1974). Prediction and evaluation of response to selection with overlapping generations. *Animal Production* **18,** 117-139.

James, J.W. & McBride, G. (1958). The spread of genes by natural and artificial selection in a closed poultry flock. *Journal of Genetics* **56**, 55-62.

Kimura, M. & Crow, J.F. (1963). The measurement of effective population number. *Evolution* 17, 279-288.

Meuwissen, T.H.E. (1997). Maximising the response of selection with a predefined rate of inbreeding. *Journal of Animal Science* 75, 934-940.

Meuwissen, T.H.E. (1998). Risk management and the definition of breeding objectives. *In: Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Vol. 25.* Armidale, pp 347-354.

Robertson, A. (1965). The interpretation of genotypic ratios in domestic animal populations. *Animal Production* 7, 319-324.

Thompson, R. (1977). The estimation of heritability with unbalanced data II. Data available on more than two generations. *Biometrics* **33**, 497-504.

Toro M. & Nieto, B. (1984). A simple method for increasing the response to artificial selection. *Genetical Research (Camb)* **44,** 347-349.

Villanueva B., Woolliams, J.A. & Gjerde, B. (1996). Optimum designs for breeding programmes under mass selection with an application in fish breeding. *Animal Science* **63,** 563-576.

Villanueva B. & Woolliams, J.A. (1997). Optimization of breeding programmes under index and constrained inbreeding. *Genetical Research (Camb)* **69,** 145-158.

Woolliams, J.A. & Mäntysaari, E.A. (1995). Genetic contributions of Finnish Ayrshire bulls over four generations. *Animal Production* **61,** 177-187.

Woolliams, J.A. & Meuwissen, T.H.E. (1993). Decision rules and variance of response in breeding schemes. *Animal Production* **56,** 179-186.

Woolliams, J.A. & Thompson, R. (1994). A theory of genetic contributions. *In: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Vol. 19.* Guelph, pp 127-134.

Wray, N.R. & Goddard, M.E. (1994). Increasing long-term response to selection. *Genetics, Selection, Evolution* **26,** 431-451.

Wray, N.R. & Thompson, R. (1990). Prediction of rates of inbreeding in selected populations. *Genetical Research (Camb)* **55,** 41-54.

Appendix A. A proof that $\mathbf{c}_t^T \mathbf{A}_t^* \mathbf{c}_t$ can be used to constrain ΔF

Wray and Thompson (1990) show that the asymptotic rate of inbreeding is approximately ¼ $\mathbf{r}^T\mathbf{r}$, where $\mathbf{r}$ is the long-term contribution vector of individuals in generation one. An examination of the proof reveals that this result is independent of the structure of the base. The proof is derived from the relationship that ½ $\mathbf{r}$ = $\mathbf{D}_1\mathbf{R}_{1,t}^T\mathbf{1}$, as $t \to \infty$, where $\mathbf{D}_1$ is ½$\mathbf{I}$, $\mathbf{R}_{j,t} = \prod_{k=j}^{t-1} \mathbf{Z}_k$ is the matrix of contributions relating descendants at generation t to ancestors from generation j ($j < t$) with the $\mathbf{Z}$ matrix defined as in Materials and Methods and $\mathbf{1}$ is a column vector with elements of one. ΔF is assumed to be a property of the breeding scheme and as such is invariant to the precise choice of the base generation i.e. which individuals are assumed to be unrelated. Hence any generation can be considered as a base and all generations (other than 0) can thus be considered as being one generation removed from the base. Therefore, the $\mathbf{r}^T\mathbf{r}$ of any generation can be used for the estimation of ΔF if the $\mathbb{D}_t$ component in $\mathbb{A}$ is set to ½$\mathbf{I}$ each generation.

The total contribution to generation t+1 of ancestors born at generation j is $\mathbf{r}_{j,t+1}$ = $\mathbf{c}_t^T\mathbb{R}_{j,t}$, where $\mathbf{c}_t$ is the vector of mating proportions at generation t. Following Wray and Thompson (1990) and substituting $\mathbb{A}^*$ for $\mathbb{A}$, the squared contributions for candidates in generation t can be expressed as:

$$\mathbf{c}_t^T \mathbb{A}^* \mathbf{c}_t = \mathbf{c}_t^T \left(\tfrac{1}{2}\mathbb{R}_{t,t}\mathbb{R}_{t,t}^T + \tfrac{1}{2}\mathbb{R}_{t-1,t}\mathbb{R}_{t-1,t}^T + \ldots + \mathbb{R}_{0,t}\mathbb{R}_{0,t}^T \right) \mathbf{c}_t$$

$$= \tfrac{1}{2} \sum_{j=0}^{t} \mathbf{c}_t^T \mathbb{R}_{j,t}^T \mathbb{R}_{j,t}^T \mathbf{c}_t + \tfrac{1}{2} \mathbf{c}_t^T \mathbb{R}_{0,t}^T \mathbb{R}_{0,t}^T \mathbf{c}_t$$

$$= \tfrac{1}{2}\sum_{j=0}^{t} \mathbf{r}_{j,t+1}^{\mathrm{T}}\mathbf{r}_{j,t+1} + \tfrac{1}{2}\mathbf{r}_{0,t+1}^{\mathrm{T}}\mathbf{r}_{0,t+1}$$

Note that $\mathbf{R}_{t,t}\mathbf{R}_{t,t}^{\mathrm{T}} = \mathbf{I}$ and the $\mathbf{R}_{0,t}\mathbf{R}_{0,t}^{\mathrm{T}}$ has been split into two terms. Similarly for generation t+1

$$\mathbf{c}_{t+1}^{\mathrm{T}}\mathbf{A}^{*}\mathbf{c}_{t+1} = \tfrac{1}{2}\sum_{j=0}^{t+1} \mathbf{r}_{j,t+2}^{\mathrm{T}}\mathbf{r}_{j,t+2} + \tfrac{1}{2}\mathbf{r}_{0,t+2}^{\mathrm{T}}\mathbf{r}_{0,t+2}$$

Let $\Delta C = \mathbf{c}_{t+1}^{\mathrm{T}}\mathbf{A}^{*}\mathbf{c}_{t+1} - \mathbf{c}_{t}^{\mathrm{T}}\mathbf{A}^{*}\mathbf{c}_{t}$ (which is held constant over generations), $\Delta r_j^2(t+1) = \mathbf{r}_{j,t+2}^{\mathrm{T}}\mathbf{r}_{j,t+2} - \mathbf{r}_{j,t+1}^{\mathrm{T}}\mathbf{r}_{j,t+1}$ and $\Delta r_{t+1}^2(t+1) = \mathbf{r}_{t+1,t+2}^{\mathrm{T}}\mathbf{r}_{t+1,t+2} = \mathbf{c}_{t+1}^{\mathrm{T}}\mathbf{c}_{t+1}$, then the increment in the constraint can be considered as a set of increments in the squared contributions:

$$\Delta C = \tfrac{1}{2}\sum_{j=0}^{t+1} \Delta r_j^2(t+1) + \tfrac{1}{2}\Delta r_0^2(t+1) \qquad 1$$

Under near equilibrium conditions which hold over the period of convergence, the contributions converge ($\Delta r_j^2(t) \to 0$ as t increases) at an approximately constant rate and hence $\Delta r_j^2(t) \approx \Delta r_{j+1}^2(t+1) = \Delta_{t-j}$ i.e. each of the increments in contributions from an ancestral generation to successive generations of descendants depends only on the number of generations that separate the ancestral generation from their descendants. Convergence implies that (i) Δ_{t-j} decreases to zero as t–j increases and (ii) $\sum_{t-j=s}^{\infty}\Delta_{t-j}$ tends to zero as t–j increases for sufficiently large s. Hence $\Delta C = \tfrac{1}{2}\sum_{k=0}^{t+1}\Delta_k + \tfrac{1}{2}\Delta_{t+1}$, which converges to half the total increment in squared contributions for a single generation i.e. from 1

$$\Delta C = \tfrac{1}{2}\left(\Delta r_{t+1}^2(t+1) + \ldots + \Delta r_0^2(t+1)\right) + \tfrac{1}{2}\Delta r_0^2(t+1)$$

$$\approx \tfrac{1}{2}\left(\Delta r_t^2(t) + \Delta r_t^2(t) + \ldots + \Delta r_t^2(t+s)\right) = \tfrac{1}{2}\mathbf{r}_{t,\infty}{}^{\mathrm{T}}\mathbf{r}_{t,\infty}$$

where s is sufficiently large and $\mathbf{r}_{t,\infty}$ is the long-term contribution vector of individuals in generation t. Therefore from Wray and Thompson (1990), $\Delta C \approx 2\Delta F$. This equivalence is refined in Appendix B.

Appendix B. Calculation of ΔC

Following Woolliams and Thompson (1994) the asymptotic rate of inbreeding can be expressed as

$$\Delta F = (1-\alpha)\, X_1 (2-X_0)^{-1}$$

where α is the extent of non-random mating as defined by Kimura and Crow (1963) and X_0 and X_1 are the mean of the diagonal elements of the contribution matrices (Wray and Thompson, 1990) of the base and first generation, respectively. The rate of inbreeding can also be expressed as a function of the increment applied to the constraint (ΔC) by using the following relationships:

$\alpha = -\frac{1}{2}\Delta C$ since $\alpha \approx -\Delta F$ for random mating (Robertson, 1965)

$X_1 = \Delta C$ (Appendix A)

$X_0 = 2\Delta C$ since X_0 refers to the base generation and the base generation has twice the contribution of any other generation (Wray and Thompson, 1990).

Thus

$$\Delta F = (1+ \Delta C /2)\, \Delta C\, [2(1- \Delta C)]^{-1}$$

which leads to

$$\Delta C = 2\Delta F\, [1-3\Delta F+12(\Delta F)^2]$$

Appendix C. A proof that the constrained maximisation of genetic response using estimated breeding values of the current generation is identical to the simultaneous constrained maximisation of genetic contributions using Mendelian sampling terms from the current and all ancestral generations

Woolliams and Thompson (1994) decomposed the estimated breeding value of an individual into the weighted sum of estimated Mendelian sampling terms of itself plus all its ancestors. Given that these ancestors are not all distinct the coefficients relating the descendant and a particular ancestor can be defined as the genetic contributions of those ancestors. Thus at generation t, the vector of estimated breeding values of the current generation ($\mathbf{g}_t$) is

$$\mathbf{g}_t = \mathbf{a}_t + \sum_{j=0}^{t-1} \mathbf{R}_{j,t}\mathbf{a}_j$$

where $\mathbf{a}_t$, and $\mathbf{a}_j$ are the vectors of estimated Mendelian sampling terms of the candidates and the ancestors born in generation j, respectively and $\mathbf{R}_{j,t}$ is defined in Appendix A. Furthermore, the total contribution in generation t+1 from ancestors born at generation j can be expressed as $\mathbf{r}_{j,t+1} = \mathbf{c}_t^{\mathrm{T}}\mathbf{R}_{j,t}$, where $\mathbf{c}_t$ is the vector of mating proportions at generation t. Thus $\mathbf{c}^{\mathrm{T}}\mathbf{g} = \sum_{j=0}^{t} \mathbf{r}_{j,t+1}^{\mathrm{T}}\mathbf{a}_j$.

Following Appendix A the squared contributions for candidates in any generation can be expressed as:

$$\mathbf{c}^{\mathrm{T}}\mathbb{A}^{*}\mathbf{c} = \tfrac{1}{2}\sum_{j=0}^{t} \mathbb{r}_{j,t+1}^{\mathrm{T}}\mathbb{r}_{j,t+1} + \tfrac{1}{2}\mathbb{r}_{0,t+1}^{\mathrm{T}}\mathbb{r}_{0,t+1}$$

After several generations $\tfrac{1}{2}\mathbb{r}_{j,t+1}^{\mathrm{T}}\mathbb{r}_{j,t+1}$ becomes constant and increments in $\tfrac{1}{2}\sum\mathbb{r}_{j,t+1}^{T}\mathbb{r}_{j,t+1}$ tend to $\tfrac{1}{2}\mathbb{r}_{j,\infty}^{T}\mathbb{r}_{j,\infty}$ for any generation j, and $\tfrac{1}{2}\mathbb{r}_{0,t+1}^{\mathrm{T}}\mathbb{r}_{0,t+1}$ also tends to a constant. Hence

after convergence of the base generation contributions, maximising $\mathbf{c}^T\mathbf{g} - \lambda\mathbf{c}^T\mathbf{A}^*\mathbf{c}$ is equivalent to maximising

$$\sum_{j=0}^{t} \mathbf{r}_{j,t+1}^T \mathbf{a}_j - \tfrac{1}{2}\lambda \sum_{j=0}^{t} \mathbf{r}_{j,t+1}^T \mathbf{r}_{j,t+1} = \sum_{j=0}^{t} \left(\mathbf{r}_{j,t+1}^T \mathbf{a}_j - \tfrac{1}{2}\lambda \mathbf{r}_{j,t+1}^T \mathbf{r}_{j,t+1} \right)$$

Thus the problem is attempting to maximise the same function simultaneously for multiple generations. The opportunity for doing so increases as j increases since the contributions converge over time and these earlier generations contribute very little to the variation in $\mathbf{c}^T\mathbf{g}$ for feasible $\mathbf{c}$. There is no separate constraint on each generation, only the aggregate rate of inbreeding is subject to constraint.

In the linear function the Lagrangian multiplier λ is the weighting factor between the two components for all generations. When a fixed overall rate of inbreeding is imposed there is a slight departure from this simultaneous optimisation, since the value of λ required for each generation in isolation to achieve the target rate of inbreeding is related to the standard deviation of the estimated breeding values which slowly changes with inbreeding (unpublished results). However, the importance of this change is limited since the variation among possible solutions at any given time arises from the small number of generations for which contributions are converging rapidly and λ will vary very little for those generations. It may therefore be concluded that the procedure has only a small departure from optimality in the problem of maximising gain with constrained inbreeding.

Appendix D. An empirical upper bound for response

Assume a population is propagated from ½N_c males and ½N_c females [$N_c = (\mathbf{c}^T\mathbf{c})^{-1}$] and with each parent having a family size $S = T\ (½N_c)^{-1}$, where T is the number of candidates. The selection proportion assuming an infinite population size is $(T\mathbf{c}^T\mathbf{c})^{-1}$.

In order to maintain individual contributions over generations two offspring from each family are selected to replace each parent. The selection intensity will be lower than that assuming optimality between and within family selection and this would imply a loss in response. An estimate of this loss scaled by the accuracy of selection is $i_w(i_p)^{-1}$ where i_w is the within family selection intensity, i_p is the proportionate selection intensity, and ρ is the accuracy of the Mendelian sampling terms when the contributions converge. Let N_r be the numbers of males still represented at convergence [$(\mathbf{r}^T\mathbf{r})^{-1}$]. Some of the initial loss associated with within family selection can be recovered by the thinning out over generations of the numbers of individuals before the fixation of their contributions. Let i_b be the increase in intensity associated with the change in the numbers represented from N_c and N_r. This thinning out is more pronounced with lower heritability since the increase in accuracy of the Mendelian sampling terms ($\delta\rho$) over time is higher and so there is greater opportunity for reducing the numbers selected.

Thus the expected upper bound to response (ΔG_{ub}) accounting for both accuracy and dispersion of genes can be expressed as

$$\Delta G_{ub} = \left(i_w\rho + i_b\delta\rho\right)\left(i_p\right)^{-1}\ \Delta G_{ideal}$$

where ΔG_{ideal} is the idealised response.

Table 1. Rates of inbreeding (ΔF) per generation (t), sum of squared contributions ($\mathbf{r}^T\mathbf{r}$), effective numbers of parents of each sex (N), and rates of genetic gain (ΔG, phenotypic standard deviation units) when using Algorithm I (with standard A), Algorithm II (with standard A and modified constraint), or Algorithm III (with the augmented A and modified constraint). The rate of inbreeding per generation was constrained to 0.025 [*]

	Algorithm I				Algorithm II				Algorithm III			
t	ΔF	$\mathbf{r}^T\mathbf{r}/4$[†]	N	ΔG	ΔF	$\mathbf{r}^T\mathbf{r}/4$	N	ΔG	ΔF	$\mathbf{r}^T\mathbf{r}/4$	N	ΔG
0	-	0.024	10.1	-	-	0.021	10.1	-	-	0.021	10.9	-
4	0.026	0.025	8.8	0.310	0.023	0.021	9.9	0.299	0.023	0.023	9.6	0.303
8	0.030	0.028	8.1	0.295	0.023	0.020	10.5	0.272	0.025	0.023	9.6	0.291
12	0.033	0.032	7.3	0.273	0.020	0.018	11.4	0.244	0.024	0.023	9.7	0.268
16	0.036	0.036	6.7	0.253	0.019	0.017	12.1	0.216	0.024	0.023	9.5	0.242
20	0.044	-	6.0	0.241	0.018	-	12.9	0.209	0.025	-	9.8	0.220

[*] Standard errors were 0.011 for ΔG, 0.001 for ΔF and 0.2 for N.

[†] The long-term contributions were calculated as the contributions from individuals born in generation t to generation t+5.

Table 2. The regression coefficient of long term contributions on mating proportions for different constraints on the rate of inbreeding (ΔF) and heritabilities (h^2)*

	ΔF		
h^2	0.05	0.025	0.00625
0.01	0.93	0.96	0.98
0.25	0.94	0.97	0.98
0.99	1.05	1.04	1.02

* Standard errors ranged from 0.01 to 0.05.

Table 3. The ratio of the sum of squared contributions to the sum of squared mating proportions for different constraints on the rate of inbreeding (ΔF) and heritabilities (h^2)*

	ΔF		
h^2	0.05	0.025	0.00625
0.01	2.33	1.89	1.41
0.25	1.92	1.72	1.35
0.99	1.30	1.27	1.19

* Standard errors were less than 0.0002 for the sum of the squared long-term contributions and 0.0004 for the sum of the squared mating proportions.

Table 4. The efficiency (%) of the observed genetic response relative to the predicted ideal ($\Delta G_{obs}/\Delta G_{ideal}$) for different constraints on the rate of inbreeding (ΔF) and heritabilities (h^2). ΔG_{obs} was calculated as the sum of products of the long-term contributions and Mendelian sampling terms of individuals born at generation three *

	ΔF		
h^2	0.05	0.025	0.00625
0.01	10.6	13.4	10.8
0.25	45.7	44.5	44.2
0.99	87.4	87.3	81.4

* Standard errors for ΔG_{obs} ranged from 0.05 to 0.005

Table 5. Asymptotic genetic responses obtained with Algorithm III, constrained sib-index and mass selection for different constraints on the rate of inbreeding (ΔF) and heritabilities (h^2)* . The responses sib-index and mass selection are expressed as a proportion of those with Algorithm III.

		ΔF		
Selection	h^2	0.05	0.025	0.00625
Algorithm III	0.01	0.019	0.015	0.009
	0.25	0.350	0.309	0.167
	0.99	1.772	1.147	0.633
sib-index	0.01	1.00	1.00	0.77
	0.25	0.93	0.87	0.81
	0.99	0.71	0.91	0.83
mass	0.01	0.79	0.87	0.77
	0.25	0.91	0.84	0.81
	0.99	0.65	0.84	0.73

* Standard errors for ΔG obtained with Algorithm III ranged from 0.05 to 0.005

Paper 34

The relationship between rates of inbreeding and expected genetic contributions

by

J.A. Woolliams
Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, U.K.

Tel: +131 527 4200
Fax: +131 440 0434
E-mail: john.woolliams@bbsrc.ac.uk

Short title: Inbreeding and expected contributions

INTRODUCTION

Wray and Thompson (1990) showed a fundamental relationship between the expected squared genetic contribution and rates of inbreeding for random mating in discrete generations. One impact of this relationship was that it tied rates of inbreeding to the numerator relationship matrix for the first time, thus narrowing the conceptual gap between the tool used in genetic evaluation of individuals and one of the key properties of a breeding scheme. Another important impact was to set out in a formal way a model for the mechanics of inheritance of selected advantage, a concept that Robertson had introduced but left unclarified. An achievement of the methods of Wray and Thompson was to obtain, for the first time, accurate predictions of ΔF in mass selection through modelling pathway extensions. However, this was done by using a recursive algorithm, so that although the mechanics were clear, the overall structure of the prediction remained obscure.

Woolliams, Wray and Thompson (1993) advanced the understanding of the structure of the prediction using expected long-term contributions. A closed form for the prediction of ΔF was derived, and it was shown to have terms involving variances of family size in 1 generation, with additional terms for the expected proliferation or diminution of ancestral lines over many generations. Furthermore it was clear that under equilibrium conditions, the model would lend itself to geometric summation across generations. Thus there were simple forms for the expected long-term contribution of an ancestor. Wray, Woolliams and Thompson (1994) extended the methods to index selection, although the form of the model is a hybrid of the approach of Woolliams, Wray and Thompson (1993) and Hill (1994). The conditional arguments of pathway extension carried out in a step-wise fashion, as conducted by Wray, Woolliams and Thompson were found to be very complex. Nevertheless worthwhile predictions were made available in a tractable form.

Santiago, Caballero and Hill (1995) used an approach that made no reference to the theory of contributions to predict ΔF in mass selection. They obtained a neater form for ΔF through a recursive argument on total drift relating the change through selection to loss of genetic variance. Unlike earlier work, these predictions were based upon equilibrium genetic variance.

Woolliams and Thompson (1994) indicated that changing the conditioning arguments employed using contributions would alter the form of predictions, for example, conditioning upon an individual's mates. This was taken further by Nagai, who showed that using this additional conditioning, the approaches adopted for mass selection by Woolliams, Wray and Thompson and Santiago, Caballero and Hill may be reconciled. Nagai has also extended the work on mass selection to overlapping generations.

This paper, therefore, examines the issues raised by the work described above. It considers the necessity for, and the role of, an unselected base generation in deriving the relationship between ΔF and the expected squared contributions, and it extends the scope of the proof to include non-random mating. It re-examines the predictions of ΔF using the equilibrium methods developed by Woolliams, Bijma and Villanueva (1998) to encompass index selection and overlapping generations and variable family size, and examines the extension of the predictions to BLUP selection and mating designs.

MATERIALS AND METHODS

The gene frequency at time t in the parents of sex x, for a unique allele at a neutral locus (say allele M, in individual j^*) in the base population can be described in terms of genetic contributions in a similar way to equation (1) of Woolliams, Bijma and Villanueva (1998). Let A_j be the gene frequency of an allele M in individual j where $A_j = 1$, ½ or 0 if j is MM, $M\bullet$, or $\bullet\bullet$ respectively, then:

$$P_t = \sum_j r_{j,0}(t) A_{j,0} + \sum_{s=1}^{t} \sum_j r_{j,s}(t) a_{j,s} \qquad (1)$$

where $r_{j,s}(t)$ is the genetic contribution of individual j born at time s, with breeding value for frequency $A_{j,s}$ and $a_{j,s} = A_{j,s} - \frac{1}{2}(A_{\text{sire}} + A_{\text{dam}})$. The $a_{j,s}$ are Mendelian sampling terms, but the variance of the sampling will depend on A_{sire} and A_{dam}; $\text{Var}(a_{j,s}) = 0$ if both A_{sire} and A_{dam} are homozygotes, ⅛ if they are both heterozygotes, or 1/16 otherwise. Generation 0 is a base generation of an unspecified kind, and since M is unique, $A_{j,0}$ is 0 for all individuals except for $A_{j^*,0}=\frac{1}{2}$. Here P_t is the average of the gene frequency in either male or female parents at generation t, and so $r_{j,t}(t) = X^{-1}$ where X is the number of parents of the same sex as j.

Random mating. F_t can be defined explicitly for any generation in the form:

$$F_t = \sum_{alleles} \sum_{j} r_{j,0}(m,t-1) r_{j,0}(f,t-1) A_{j,0}^2 + \sum_{alleles} \sum_{s=1}^{t-1} \sum_{j} r_{j,s}(m,t-1) r_{j,s}(f,t-1) a_{j,s}^2$$

where $r_{j,s}(x,t\text{-}1)$ is the average contribution to parents of sex x at time t-1 . This is obtained from the product of the gene frequencies in the male and in female parents. The cross-product terms in $A_{j,0} A_{k,0}$ are zero since for all alleles A=0 except for just a single individual, and the Mendelian sampling terms from different individuals are independent of all other terms for a neutral locus. More precisely for each allele and each ancestor, the term $r_{j,s}(m,t-1) r_{j,s}(f,t-1) a_{j,s}^2$ should be the sum of products of contributions of the ancestor to each male and female mating pair i.e. $\sum_{mates\ (i(m),i(f))} r_{j,s}(i(m),t-1) r_{j,s}(i(f),t-1) a_{j,s}^2$ but with random mating this can be replaced by the expectation which is the product of the average contribution of the ancestor to the male parents and to the female parents.

This can be simplified by noting $\sum_{alleles} E[a_{j,s}^2] = \frac{1}{4}$ for $s = 1$ and $\frac{1}{4}(1-\Delta F)^{s-2}$ for $s \geq 2$ and the sum over generation 0 is $\frac{1}{2}\sum_{j} r_{j,0}(m,t-1) r_{j,0}(f,t-1)$ (see Appendix 1). The neutrality of the locus is again critical since otherwise a covariance of r_j and a_j will arise (Woolliams and Thompson, 1994). The summation over cross products $\sum_{j} r_{j,s}(m,t-1) r_{j,s}(f,t-1)$ will be written $C_s(t\text{-}1)$. Note $C_{t-1}(t-1) = 0$ since an individual without offspring cannot contribute to both sexes. Therefore:

$$F_t = \tfrac{1}{2}C_0(t-1) + \tfrac{1}{4}C_1(t-1) + \tfrac{1}{4}C_2(t-1) \ldots + \tfrac{1}{4}(1-\Delta F)^{t-4}C_{t-2}(t-1)$$

$$F_{t+1} = \tfrac{1}{2}C_0(t) + \tfrac{1}{4}C_1(t) + \tfrac{1}{4}C_2(t) \ldots + \tfrac{1}{4}(1-\Delta F)^{t-4}C_{t-2}(t) + \tfrac{1}{4}(1-\Delta F)^{t-3}C_{t-1}(t)$$

$$F_{t+1}-F_t = \tfrac{1}{2}C_0(t) + \tfrac{1}{4}C_1(t) + \tfrac{1}{4}C_2(t) \ldots + \tfrac{1}{4}(1-\Delta F)^{t-4}C_{t-2}(t) + \tfrac{1}{4}(1-\Delta F)^{t-3}C_{t-1}(t)$$
$$-\tfrac{1}{2}C_0(t-1) - \tfrac{1}{4}C_1(t-1) - \tfrac{1}{4}C_2(t-1) \ldots -\tfrac{1}{4}(1-\Delta F)^{t-4}C_{t-2}(t-1)$$

Assuming an equilibrium then a steady state of pedigree development will occur and $E[C_j(k)] = E[C_{j-1}(k-1)]$. This is not a strong assumption in the context of the problem since in the absence of an equilibrium there would be no single ΔF to predict. Therefore the terms in $C_j(t)$ can be reduced to terms in $C_{j-1}(t-1)$.

$$F_{t+1}-F_t = -\Delta F(\tfrac{1}{2}C_0(t-1)+\tfrac{1}{4}C_1(t-1)+\ldots+\tfrac{1}{4}(1-\Delta F)^{t-2}C_{t-2}) + \tfrac{1}{2}C_0(t)-\tfrac{1}{4}C_0(t-1)+\tfrac{1}{2}\Delta F C_0(t)$$

The last term, $\tfrac{1}{2}\Delta F\ C_0(t)$, is required to be added back in order to balance the terms. The initial bracketed terms are equal to $-(\Delta F)F_t$. The assumption of panmixia gives $C_0(t) - C_0(t-1) \to 0$ and differences in the contributions to males and to females from an ancestor also disappear as $t \to \infty$. Therefore $C_0(t) \to \sum_j E[r_j^2]$ and

$$\Delta F = \tfrac{1}{4}\sum_j E[r_j^2]\ (1+2\Delta F) \approx \tfrac{1}{4}\sum_j E[r_j^2]$$

This derivation differs from Wray and Thompson (1990) in that it does not subtract out the base population and ΔF is defined in terms of the base. In their derivation the base was unselected and therefore not in equilibrium at the start of the selection process which encouraged its special treatment. The choice of generation on which the estimate is obtained is arbitrary except that it is at the start of some local equilibrium in which some 'equilibrium ΔF' may exist. The correction $2\Delta F$ (Woolliams, Wray and Thompson, 1993) arises naturally in this proof. The derivation using the explicit formula of inbreeding for an assumed allele is also preferable.

The sum is over the selected parents, $\Delta F \approx \tfrac{1}{4}(ME[r_{i(m)}^2] + FE[r_{i(f)}^2])$ and where expectations are conditional on selection. Since for unselected candidate $r_{i(x)} = 0$, this can be re-expressed as $\Delta F \approx \tfrac{1}{4}(T_m E[r_{i(m)}^2] + T_f E[r_{i(f)}^2])$ where the expectations are now unconditional.

Non-random mating. Here non-random mating is defined by departures from Hardy-Weinberg equilibrium. This departure will be measured by α which is the correlation between uniting gametes for a neutral allele. Robertson showed that even when male and female mates were randomly allocated, α is negative and of the order of the rate of inbreeding. Two changes need to be considered in developing the proof for random mating, one in the Mendelian variance and secondly in the definition of $C_j(t)$. The Mendelian sampling variance for generation 1 remains ¼, irrespective of any established pedigree structure in the base, because the alleles are re-defined in the base as being distinct. For subsequent generations, the variance is $¼(1-\alpha)$.

The second change is that $C_j(t)$ no longer be the simple sum of cross-products of the average contributions of the ancestor to the current male and female parents. Since the mating is non-random the more precise definition, given earlier, must be used. However this re-definition is not a complication when the assumption of equilibrium is made. The proof for random mating is carried through as before and the only terms that remain are those for $C_0(t)$ and $C_2(t)$ for large t. In a panmictic population for large t the contributions of an ancestor in early generations will be equal for all individuals in generation t (but different from other ancestors). Thus $C_j(t)$ will, as before, tend to $\Sigma E[r_j^2]$ irrespective of the non-random mating. Therefore the following is obtained:

$$\Delta F = ¼C_0(t) - ¼\alpha\ C_2(t) + ½\Delta F\ C_0(t)$$

Assuming equilibrium $\Delta F = ¼(ME[r^2_{i(m)}] + FE[r^2_{i(f)}])(1-\alpha+2\Delta F)$ which may be approximated in the same way as for random mating by $\Delta F = ¼(1-\alpha)(ME[r^2_{i(m)}] + FE[r^2_{i(f)}])$. This result was given by Woolliams and Thompson (1994).

Overlapping generations. If ΔF is taken per cohort then the structure of the preceding proof holds and

$$\Delta F \text{ per cohort} = ¼(1-\alpha+2\Delta F)(T_m E(r^2_{i(m)}) + T_f E(r^2_{i(f)}))$$

where the expectation is over the squared contributions from a single cohort, and the expectations are not conditional on selection. The correction factor for overlapping generations is ½ΔF over all cohorts that constitute a generation, thus it is 2 × (ΔF per generation).

RELATIONSHIP BETWEEN ΔF AND EXPECTED CONTRIBUTIONS

The objective of the following section is to develop a relationship between ΔF and expected contributions (μ). Since $\Delta F \propto E[r^2] = \mu^2 + \sigma^2$, such a result is obtained by relating the variance of the contribution σ^2 to its expectation. This will be done by first considering the form of σ^2 conditional on a full set of variables $s_{i(x)}$ that confer selective advantage on the offspring, and then as a second step looking at the unconditional expectation.

Wray and Thompson (1990) give an analogy which will act as a starting point. Consider the variance of a variable v which depends on variable n, then $Var(\nu) = E_n[Var(\nu|n)] + Var_n(E(\nu|n))$. If v occurs subsequently to n, and is a sum of n independent random variables $s_1 \ldots s_n$ with mean μ_s and variance σ_s^2, then $E[\nu] = \mu_n\mu_s$ and $Var(\nu) = \mu_n\sigma_s^2 + \sigma_n^2\mu_s^2$.

Whilst the extension to two sexes made by Wray and Thompson (1990) is relevant, their application of the analogy is made back to front. Thus they use it to move from summing over generations 1 to n to summing over generations 1 to (n+1) by adding terms for (n+1). Thus ΔF arises after (in principle) an infinity of recursive steps (although few were necessary for convergence). However, in all these steps (co)variance matrices and difficult path extension arguments were required.

If an equilibrium is assumed, then the extension can be made by assuming 2 to n+1 is the same as 1 to n, and so modelling the extension of 1 to 2 is sufficient for to obtain 1 to n+1 and, consequently, for all steps. The special case of the long-term is able to be carried out in one step since the extension by one generation must yield the same result.

The following shall initially assume random mating and discrete generations. For an ancestor $i(x)$ conditional on a vector of full information on selective advantages $\boldsymbol{s}_{i(x)}$ then:

$$r_{i(x)}|\boldsymbol{s}_{i(x)} = \mu_{i(x)} + \epsilon_{i(x)} = \tfrac{1}{2}\sum_{y=m,f} (E[n_y|\boldsymbol{s}_{i(x)}]+e_{i(x),y}) (E[r_{j(y)}|\boldsymbol{s}_{i(x)}] + \epsilon'_{j(y)})$$

where $\mu_{i(x)}$ is $E[r_{i(x)}|\boldsymbol{s}_{i(x)}]$, n_y is the number of offspring selected of sex y and $j(y)$ is a selected offspring of sex y. Since the conditioning is on a *full set of selective advantages* (including common environment and litter effects) the variance of the errors for the number selected will be binomial within any litter.

The analogy of Wray and Thompson is made if the s variables are identified with $(r_{j(y)}|\boldsymbol{s}_{i(x)})$ and the n variables identified with $(n_y|\boldsymbol{s}_{i(x)})$. Let $Var(\epsilon'_{j(y)}) = \nu_y^2$ and $E[r_{j(y)}|\boldsymbol{s}_{i(x)}] = \mu^*_{i(x),y}$ will be assumed to depend on the sex of the offspring y; $\boldsymbol{\mu}_{i(x)}^* = (\mu_{i(x),m}^*, \mu_{i(x),f}^*)^T$. The variables $(n_m, n_f |\boldsymbol{s}_{i(x)})$ have a bivariate distribution with expectation $\boldsymbol{\theta}_{i(x)}$ with (co)variance matrix $\boldsymbol{V}_n$.

Assumption 1. $\boldsymbol{V}_n$ will be assumed independent of $i_{(x)}$ and this may be justified via the following analogy. For a binomial with n trials with probability of success p_i, where p_i has mean p and variance σ_p^2 then the variance of the number of successes is $np(1-p) + n(n-1)\sigma_p^2$; the terms in σ_p^2 are n^2, arising from variation in the expected number of success for a given p_i, and $-n$ from the mean of the variance for a given p_i. Therefore, the dominant term in σ_p^2 arises from modelling the dependence of the expectation of the number of successes for each ancestor rather than the dependence of the variance on each ancestor.

Assumption 2. If litter sizes are Poisson of size n, then binomial trials with probability p has a variance np. The sex of each individual within the litter is independent of the sex of the littermates. Combined with the first assumption, $\boldsymbol{V}_n$ for each litter of $i(x)$ has the form $(MX^{-1}Z^{-1}, 0|0\ FX^{-1}Z^{-1})$ where Z is the number of litters the ancestor has. Sets of full sibs may be taken from separate litters, with any litter effects accounted for in the conditioning vector

$\boldsymbol{s}_{i(x)}$.

Therefore, using the conditioning analogy of Wray and Thompson (1990):

$$E[\sigma^2_{i(x)}|\boldsymbol{s}_{i(x)}] = \tfrac{1}{4}(\nu^2_m,\nu^2_f)^T\theta_{i(x)} + \tfrac{1}{4}\sum_{litters}\boldsymbol{\mu}^*_{i(x)}\boldsymbol{V}_n\boldsymbol{\mu}^*_{i(x)}.$$

$\theta_{i(x)}$ will depend on $\boldsymbol{s}_{i(x)}$, however, since ν^2_m and ν^2_f are assumed independent of i and $\boldsymbol{s}_{i(x)}$ and $\sum_i \boldsymbol{s}_{i(x)} = 0$ there is no net contribution to $\sum_i \sigma^2_{i(x)}$ from $\boldsymbol{s}_{i(x)}$. Thus the first of the terms on the right hand side may be simplified to $\tfrac{1}{4}(\nu^2_m MX^{-1} + \nu^2_f FX^{-1})$, since MX^{-1} and FX^{-1} are the average number of offspring of each sex per parent of sex x. (It is also reasonable therefore to assume σ^2_x has no terms linear in $z_{i(x)}$.) The second term also simplifies to $\tfrac{1}{4}(\mu^{*2}_{i(x),j(m)}MX^{-1} + \mu^{*2}_{i(x),j(f)}FX^{-1})$.

$$E[\sigma^2_x|\boldsymbol{s}_{i(x)}] = \tfrac{1}{4}(\mu^{*2}_{i(x),m} + \nu^2_m)MX^{-1} + \tfrac{1}{4}(\mu^{*2}_{i(x),f} + \nu^2_f)FX^{-1}$$

Note that M and F are simply the numbers of parents and not assumed to arise from a hierarchical mating design.

The second step is to take expectations over $\boldsymbol{s}$. With the equilibrium condition $E_s[\mu^{*2}_{i(x),j(y)}+\nu^2_y] = E_s[\mu^2_{i(y)}] + \sigma^2_y$. Therefore, by substitution:

$$\sigma^2_m = \tfrac{1}{4}(E_s[\mu^2_{i(m)}] + \sigma^2_m) + \tfrac{1}{4}FM^{-1}(E_s[\mu^2_{i(f)}] + \sigma^2_f)$$

$$\sigma^2_f = \tfrac{1}{4}MF^{-1}(E_s[\mu^2_{i(m)}] + \sigma^2_m) + \tfrac{1}{4}(E_s[\mu^2_{i(f)}] + \sigma^2_f)$$

Solving these equations, we obtain $\sigma^2_x = \tfrac{1}{2}E_s[\mu^2_{i(x)}] + \tfrac{1}{2}X'X^{-1}E_s[\mu^2_{i(x')}]$ where X' is the sex other than X. Then since $\Delta F \approx \tfrac{1}{4}ME[r^2_{i(m)}] + \tfrac{1}{4}FE[r^2_{i(f)}]$, and substituting $E[r^2] = \mu^2 + \sigma^2$ the following result is obtained.

$$\Delta F = \tfrac{1}{2}(ME_s[\mu^2_{i(m)}] + FE_s[\mu^2_{i(f)}])$$

The rate of inbreeding is therefore the average of the sum of squared expected contributions from each sex. The foregoing method has relied heavily on the removal of litter specific effects and maternal environmental effects by $\boldsymbol{s}_{i(x)}$. The result justifies this: it is easier, and simpler conceptually to use the methodology developed by Woolliams, Bijma and Villanueva (1997) to

account for these in $\mu_{i(x)}$ and to benefit from the result shown, than to try to account for them in other ways.

Overlapping generations assuming Poisson litter sizes. With overlapping generations, the progeny of an individual is not only split into two sexes, but also into subclasses representing future breeding categories. A category will depend on sex, age, breeding use (e.g. to breed females only or to breed both sexes; or to be superovulated or not). The categories in this case should also include the potential to breed at a specific combination of ages (e.g. at age 1 only or at both ages 1 and 2). Thus the number of categories will inevitably depend on the complexity of the breeding scheme; however the definition of the categories is simply one of enumeration. The methodology for obtaining expected long-term contributions is given by Woolliams, Bijma and Villanueva (1997). Unlike Woolliams, Bijma and Villanueva (1997) it is important conceptually to define the categories so that each parent appears only in one category. Let X_q be the number of individuals from a cohort that will be in category q and let $\mu_{i(q)} = \alpha_q + \beta_q^T(s_{i(q)} - \overline{s_q})$

For category q, the expected number of progeny of category p is n_{pq} and let $g^*_{pq} = \frac{1}{2}n_{pq}X_p^{-1}$ where ½ turns g^* into an expected gene flow between categories. Since there is an equilibrium, from equation (5) of Woolliams, Bijma and Villanueva (1997), $\alpha_q = \sum X_q^{-1} g^*_{pq} X_p \alpha_p$; or equivalently $(N\alpha) = (G^{*T})(N\alpha)$ where N is diagonal, with X_q on the diagonal. Also since the rows of G^* sum to 1, $1^T G^{*T} = 1^T$. The definition of G^* has been made consistent (parents in columns, offspring in rows) with gene flow as described by Hill (1974), although this means that it is the transpose that most often appears in formulae.

V_n is now diagonal with dimensions equal to the number of categories, and for a category q parent, with Poisson litter sizes, the pth diagonal element is equal to $2g^*_{pq}X_pX_q^{-1}$. Then following similar arguments to the discrete case:

$$N\sigma^2 = \tfrac{1}{2}(G^{*T})N\sigma^2 + \tfrac{1}{2}(G^{*T})N\mu^2 \qquad (5)$$

$$N\sigma^2 = (I-(\tfrac{1}{2}G^{*T}))^{-1}(\tfrac{1}{2}G^{*T})N\mu^2$$

Thus the vector of variances of the long-term contributions is once again expressed in terms of the expected squared mean. However, the contribution to the inbreeding is given by $\mathbf{1}^T N\sigma^2$. Using equation (5) and $\mathbf{1}^T G^{*T} = \mathbf{1}^T$ gives:

$$\mathbf{1}^T N\mu^2 = \mathbf{1}^T N\sigma^2$$

$$\Delta F \approx \tfrac{1}{2}\mathbf{1}^T N\mu^2 \qquad (6)$$

This expresses the rate of inbreeding per cohort as ½ times the sum of squared expected long-term (lifetime) contributions for all individuals in a cohort. Likewise ΔF per generation is L times this value i.e. defined for a generation of individuals not just for a cohort.

Litter sizes other than Poisson

The assumption of Poisson litter sizes in random selection results in the classical formulae of Wright (1932) and Hill (1977) with terms containing only the numbers of parents of each sex. Departures from Poisson are likely to be second order corrections to these. Therefore in what follows the terms in $\boldsymbol{\mu}^*$ that depend on $\boldsymbol{s}$ will be ignored; they will in any case be multiplied by π_{pq} (see Woolliams, Bijma and Villanueva, 1997) which further reduces their importance.

The form of V_n will depend on q and can be divided into $V_{n(q),P} + V_{n(q),dev}$ where the second term contains the departures caused by deviations from Poisson litter sizes. For each q and neglecting terms in s (as justified above), there will be a term δ_q defined by $\alpha^T V_{n(q),dev}\alpha$ which summarizes the effect of departures from Poisson for category q. Let δ be the vector of these departures and note that δ can be negative since it is a deviation from Poisson and is not itself a (co)variance matrix. Then:

$$N\sigma^2 = (I-(\tfrac{1}{2}G^{*T}))^{-1}(\tfrac{1}{2}G^{*T})N\mu^2 + (I-(\tfrac{1}{2}G^{*T}))^{-1}N\delta \qquad (7)$$

Since: (i) $\mathbf{1}^T G^{*T} = \mathbf{1}^T$; (ii) $(I-(\tfrac{1}{2}G^{*T}))^{-1}$ can be expanded in powers of $\tfrac{1}{2}G^{*T}$; and (iii) the

contribution to ΔF depends upon $\mathbf{1}^T N\sigma^2$; the second term in equation (7) contributes a total of $\frac{1}{8}\mathbf{1}^T N\delta$ to ΔF. A first order correction for departures from Poisson litter sizes is then:

$$\Delta F \approx \tfrac{1}{2}\mathbf{1}^T N(\mu^2 + \tfrac{1}{4}\delta) \qquad (8)$$

When correlations of indices are not through defined selective advantages.

In some indices, such as sib indices, the forces governing co-selection are not described by an identifiable selective advantage even though correlations among sibs exist. For example, selection on within-family deviation induces negative correlations among the indices of full-sibs. These arise from the Mendelian sampling terms and environmental deviations of the candidates themselves, and can be viewed as forming correlated indices with no inherited advantages. This reduces the problem to the their inclusion into δ.

Appendix 1. The expected Mendelian sampling variance in Generation 1.

The expected Mendelian sampling variance in generation 1 summed over all alleles in the founders can be calculated using the following argument. For the progeny of the carrier founder j^* of the allele the gene frequency has mean ¼, i.e. half of the gene frequency in carrier (½) plus half of that in mate (0), with $\sigma_a^2 = 1/16$. For progeny of other parents, the $\sigma_a^2 = 0$. Therefore, for a single allele, the Mendelian sampling variance is $\sigma_a^2 = 2r_{j^*,0}(1)(16)^{-1} = \frac{1}{8}r_{j^*,0}(1)$ where the 2 corrects the contribution of j^* to generation 1 to be the number of offspring of j^*. Summing over all alleles (2 per base individual), and since the sum of all contributions must be 1, the expected variance is ¼.

In generation 2, the Mendelian sampling variance is unchanged since there will be no inbreeding in the parents, but for other generations the expected variance is $\frac{1}{4}(1-\Delta F)^{s-2}$ in generation $s>1$, where ΔF is the rate of inbreeding among the parents. Also:

$$E[\sum_{alleles}\sum_{j} r_{j,0}(m,t)r_{j,0}(f,t)A_{j,0}^2] = \frac{1}{2}\sum_{j} r_{j,0}(m,t)r_{j,0}(f,t)$$

since $A_{j,0}^2$ has a value ¼ for each of its 2 alleles, and 0 otherwise.