

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 $\frac{1}{2}$ (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 $\Delta F = \frac{1}{4} \sum r_{i(x)}^2$ 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 $E(\Delta F) = \frac{1}{4} [ME(r_{i(x)}^2) + FE(r_{i(x)}^2)]$ 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:

$$E(r_{i(x)}) = (2X)^{-1} [1 + \frac{1}{2} i S_{-} A_{i(x)} \sigma_p^{-1}] = (2X)^{-1} [1 + i(1 + kh^2)^{-1} A_{i(x)} \sigma_p^{-1}] \quad \dots(1)$$

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

$$E(r_{i(x)}) = (2X)^{-1} (1 + \frac{1}{4} i S_{-} (\tau_x + \tau) A_{i(x)} \sigma_f^{-1}) \quad \dots(2)$$

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 $c = \frac{1}{2}(1 - \tau k \beta)$ 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 $\sigma_f = \sigma_p$. 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 asymmetric 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 contribution 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:

$$\Delta F = \Delta F_E + i^2 [S_{-}^2 - 1] \rho_m [(16M)^{-1} + (16F)^{-1}] + i^2 [S_{-}^2 - 1] \rho_f (8F)^{-1} + i^2 B_{-} (S_{-} - 1) [2\rho_m + (1 + MF^{-1})\rho_f] [(32M)^{-1} + (32F)^{-1}] \quad \dots(3)$$

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 $i^2 S_w^{-2}$, 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 $iS_w A_{i(x)}$ in the long-term) the additional terms required for ΔF are of the form $i^2(S_w^{-2}-1)$. 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 $\frac{1}{4}$ by $\frac{1}{4}(1-\alpha)$ 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:

$$\Delta F = \Delta F_E + i^2 [S_w^2 - 1] \rho_m [(8M)^{-1} + (8F)^{-1}] + i^2 [S_w^2 - 1] \rho_f (4F)^{-1} \quad \dots(4)$$

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_w 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):

$$E(\Delta G) = T_m E(r_{i(m)} a_{i(m)}) + T_f E(r_{i(f)} a_{i(f)}) \quad \dots(5)$$

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 $E(a_{i(x),t}^2) = 2\sigma_0^2$ 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), $E(r_{i(x)}a_{i(x)} | i \text{ selected}) = (2X)^{-1} i S_{-}(1 - k_x h^2) h_0^2 \sigma_p^{-1}$, and $E(r_{i(x)}a_{i(x)} | i \text{ not selected}) = 0$. To complete the unconditional expectation note that $E(a_{i(x)} | i \text{ selected}) = \frac{1}{2} i_j h_0^2 \sigma_p^{-1}$ and $E(r_{i(x)} | i \text{ selected}) = (2X)^{-1}$ and $Prob(i \text{ selected}) = XT_x^{-1}$; and, after summing over the sexes:

$$E(\Delta G) = \frac{1}{2} i h_0^2 \sigma_p^{-1} (1 + \frac{1}{2} S_{-}(1 - k h^2)) = i h_0^2 \sigma_p^{-1} (1 + k h^2)^{-1} \quad \dots(6)$$

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 $i h_e^2 \sigma_p$ 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 $i h_e^2 \sigma_p$ 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.

$$E(\Delta F) = \frac{1}{4} [T_m E(r_{i(m)}^2) + T_f E(r_{i(f)}^2)]$$

$$E(\Delta G) = T_m E(r_{i(m)} a_{i(m)}) + T_f E(r_{i(f)} a_{i(f)})$$

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:

$$\sum_{x=m,f} T E(\frac{1}{4} r_{i(x)}^2 - \lambda r_{i(x)} a_{i(x)}) = \frac{1}{4} T \sum_{x=m,f} E(r_{i(x)} (r_{i(x)} - 4\lambda a_{i(x)})) \quad \dots(7)$$

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 $(2X)^{-1} + \alpha_x (a_{i(x)} - G_{w(x)}) + \epsilon_{i(x)}$ 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:

$$\sum_{x=m,f} X E([(2X)^{-1} + \alpha_x (a_{i(x)} - G_{w(x)})] [(2X)^{-1} + (\alpha_x - 4\lambda) (a_{i(x)} - G_{w(x)})] - 4\lambda G_{w(x)}) \quad \dots(8)$$

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), $\lambda = \frac{1}{2} (G_{req} - G_w) [E(M a_{i(m)}^2 + F a_{i(f)}^2 | i, j \text{ selected})]^{-1}$. 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:

$$\Delta F \geq 3(32M)^{-1} + (32F)^{-1} + \frac{1}{4} (G_{req} - G_w)^2 [E(M a_{i(m)}^2 + F a_{i(f)}^2 | i, j \text{ selected})]^{-1} \quad \dots(9)$$

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 , 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 $h_g^2=1$ 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, \dots, A_n used in proportions p_1, \dots, p_n with $(\sum p_i=1)$, to maximize clonal value $(\sum p_i A_i)$ with a constraint on diversity $(\sum p_i^2)$ 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 j th 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:

$$I \propto \hat{a}_i + \sum_t \sum_{j=1}^{2^t} \hat{a}_{ij,t} (\frac{1}{2} - \lambda) 2^{-t+1}$$

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.

$$I \propto \hat{a}_i + \sum_t \sum_{j=1}^{2^t} \hat{a}_{ij} 2^{-t} c_{jt}$$

where $c_{jt}=1$ for $t < t^*$, $0 \leq c_{jt} \leq 1$ for $t = t^*$ and $c_{jt}=0$ 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 LATTEY, 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 $\frac{1}{2}$. This constraint was invoked to justify the replacement of $A_{i(x)}$ by $(A_{i(x)} - A_{bw})$ throughout (where A_{bw} 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 genealogical 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.