

PREDICTION OF EFFECTIVE SIZE OF POPULATIONS UNDER SELECTION AND NON-RANDOM MATING

A. Caballero and E. Santiago*

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland; *Departamento de Biología Funcional, Universidad de Oviedo, Oviedo 33071, Spain.

SUMMARY

A general equation to predict effective size of populations under continued selection is presented, which considers the possibility of non-random mating. This includes assortative mating, deliberate mating between full sibs and a simple system of mating (compensatory mating) which is proposed to reduce rates of inbreeding with practically no loss of response to selection. This mating system, in which individuals from the families in which many are selected are mated to others with few selected, is simulated alone and in combination with another selection method to reduce rates of inbreeding under BLUP selection.

INTRODUCTION

The concept of effective population size was introduced by Wright (1931) and its prediction has been developed since then for a number of different circumstances (see Caballero [1994] for a review). Wright derived a formula in terms of variance of family size to account for differences in contributions from parents to the next generation due to non-inherited causes. The problem of inherited causes of variation was first addressed by Robertson (1961), who introduced the idea of the accumulation of selective advantages of individuals over generations. Thus, the selective advantage of an individual is expected to be reduced by one half each generation in its descendants, and the total selective advantage over generations increases in a series $1 + 1/2 + 1/4 + 1/8 + \dots$, up to a limiting value of twice the selective advantage present in the initial generation. Robertson's prediction accounts for this cumulative effect but underpredicts effective sizes, especially for high heritability and intense selection, because the reduction of genetic variance with selection is not accounted for (Wray and Thompson, 1990; see below). In order to overcome this problem, Wray and Thompson (1990) have developed a recursive method to approximate the effective size as a function of the mean and variance of the contributions of ancestors in the first generation to descendants in the limit, and Woolliams *et al.* (1994) and Wray *et al.* (1994) have derived equations to predict these means and variances for mass and index selection, respectively, and random mating.

In this paper, an alternative method which generalises Robertson's approach is presented leading to general equations which allow predictions of effective size of populations under mass selection and non-random mating. Based on this approach, a new system of mating which avoids inbreeding with little loss of response is proposed.

PREDICTION OF EFFECTIVE SIZE

Effective size is computed from the variance of change in gene frequency or the rate of inbreeding of a neutral gene, which is assumed to be unlinked to the selected genes. Changes in gene frequency are the result of three independent processes acting each generation, random association between the neutral gene and families with a selective advantage or disadvantage, random sampling of individuals among families, and Mendelian sampling of heterozygotes. The two last processes also occur when differences in contributions from parents are due to non-inherited causes. We now consider the first process regarding

inherited variation. Let C^2 be the variance of relative selective advantages of families, i.e. the variance of the expected number of individuals contributed by a family, divided by the square of the mean (4, if population size is constant over generations). In the first generation of selection, the total variance of family size is approximately the sum of the component due to selection ($4C^2$) and the variance of family size due to sampling or non-inherited causes. The association between the neutral gene and individuals with a selective advantage is expected to persist over generations, being diluted by one half each generation due to segregation and recombination. However, the dilution of the association also depends on the reduction of genetic variance by selection (Bulmer, 1971) and the correlation between the expected selective advantages of male and female parents (r). It can be shown that the association is diluted by a proportion $G(1+r)/2$ every generation, assuming G is constant over generations, where G is the remaining proportion of genetic variance in the selected individuals. Therefore, the cumulative effect of selection can be represented as the sum of an infinite series $Q = \sum_{i=0}^{\infty} [G(1+r)/2]^i = 2/[2 - G(1+r)]$. If the reduction in variance each generation is neglected (i.e. $G = 1$) and $r = 0$, $Q = 2$, as was predicted by Robertson.

The effective size (N_e) under continued selection can then be approximated by adding a term due to selection to the variance of family size due to sampling without selection (S_k^2) in the usual formula,

$$N_e = \frac{4N}{2(1 - \alpha_I) + (S_k^2 + 4Q^2C^2)(1 + \alpha_I + 2\alpha_O)}$$

(Santiago and Caballero, 1994), where N is the number of breeding individuals (half of each sex), α_I is the deviation from Hardy-Weinberg proportions and α_O is the correlation between genes in male and female parents.

The values of G and C^2 above depend on the genetic system and type of selection. For a model of truncation selection on a normally distributed trait controlled by an infinitesimal model of gene effects, $G = 1 - kh^2$ (Bulmer, 1971), where $k = i(i - x)$, i is the selection intensity (standardised selection differential), x is the truncation point of the standardised normal distribution, and h^2 is the asymptotic heritability of the trait, where 'asymptotic' refers to the nearly steady state in which the decline in heritability is negligible. Moreover, for the same model C^2 can be approximated by $i^2\rho$ (Robertson, 1961), where ρ is the asymptotic intraclass correlation of full-sib family members.

Under random mating of selected parents $\alpha_I = -1/(N - 1)$ and $\alpha_O = r \approx 0$. For a multinomially distributed sampling variance of family size and large N , $S_k^2 = 2$, $\alpha_I \approx 0$ and $N_e = N/(1 + Q^2C^2)$ (cf. Robertson, 1961). However, when an average proportion β of full-sib matings occurs each generation, $r \approx \beta$ and $\alpha_I \approx \alpha_O \approx \beta/(4 - 3\beta)$. With phenotypic assortative mating of selected parents, r is approximately $r_P\rho$, where r_P is the phenotypic correlation among mates.

Finally, predictions can also be made for a system of mating (compensatory mating) by which effective size is increased without substantial effect on response. By mating individuals from families in which many are selected to others from families with few selected, a negative correlation is generated between the changes in gene frequency of the neutral gene due to selection in one generation and the changes due to sampling of individuals between families one generation before. These opposing correlations cancel each other every generation and, as a result, the cumulative effect of selection approximately vanishes and N_e can be approximated by using $Q^2 = 1$, the first term in the series.

When numbers of male and female parents are different, such that N_m males are mated to N_f/N_m females each, and assuming random mating for simplicity, effective size for sex s can be approximated by

$$N_{es} = \frac{4N_s}{\left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}} \right] (1 - \alpha_s) + \left[\left(\frac{S_{sm}^2}{\mu_{sm}^2} + \frac{2S_{sm, sf}}{\mu_{sm}\mu_{sf}} + \frac{S_{sf}^2}{\mu_{sf}^2} \right) + 4Q^2C_s^2 \right] (1 + \alpha_s)}$$

(Santiago and Caballero, 1994), where N_s is the number of individuals of sex s , $\alpha_s (= -1/[2N_s - 1])$ under random mating) is the deviation from Hardy-Weinberg proportions in individuals of sex s , μ_{sm} (μ_{sf}) and S_{sm}^2 (S_{sf}^2) are the mean and variance of the number of male (female) offspring contributed by parents of sex s without selection, and $S_{sm, sf}$ the corresponding covariance, $C_s^2 = [(C_{sm} + C_{sf})/2]^2$, where $C_{sm}^2 \approx i_m^2 \rho_s$ and $C_{sf}^2 \approx i_f^2 \rho_s$, i_s is the selection intensity for sex s , ρ_s is the intraclass correlation of individuals of families from parents of sex s , and Q is obtained as above with $G = (G_m + G_f)/2$ and $G_s \approx 1 - k_s h^2$, where k_s is as defined above but applied to individuals of sex s . Finally, N_{es} for each sex can be combined as $N_e = 4N_{em}N_{ef}/(N_{em} + N_{ef})$.

SIMULATIONS AND DISCUSSION

Stochastic simulation was carried out to check the prediction equations. Artificial selection was made on a trait controlled by an infinitesimal model. Effective sizes were obtained from the average rate of inbreeding generally between generations 5 and 14.

The following table shows simulated (*sim*) (standard errors about 0.1) and predicted (*pre*) values of effective size for a population with 20 male and 20 female parents (except in the scheme HM), 6 individuals of each sex scored per family, and a range of initial heritabilities (h^2). Mass selection was performed in all cases and the selected individuals were mated according to the following schemes: *R*: random mating of selected parents; *N(P)*: maximum negative (positive) assortative mating; *CM*: compensatory mating, families were arranged according to the total number of individuals selected, and males with the highest ranking were mated with females with the lowest, in sequence; *HM*: hierarchical mating with 20 males and 100 females, each male randomly mated to 5 females; *FS*: 50% of the matings between full-sibs, at random otherwise.

h^2	R		N		P		CM		HM		FS	
	<i>sim</i>	<i>pre</i>	<i>sim</i>	<i>pre</i>	<i>sim</i>	<i>pre</i>	<i>sim</i>	<i>pre</i>	<i>sim</i>	<i>pre</i>	<i>sim</i>	<i>pre</i>
0.1	33.7	33.2	34.2	34.0	33.1	32.7	38.6	39.1	52.6	51.9	17.4	16.5
0.2	29.3	29.2	30.4	30.3	28.0	27.6	36.1	36.1	45.9	45.1	14.9	14.5
0.4	25.9	26.0	27.7	27.8	22.5	23.3	33.2	32.3	40.2	40.0	12.9	13.0
0.8	26.2	27.2	31.2	30.5	16.8	22.6	31.0	29.8	42.7	43.5	15.7	15.5

The average absolute error in the predictions relative to the simulated values is 1.6% if the case of positive assortative mating and $h^2 = 0.8$ is not considered. This exceptional severe overprediction occurs because the approximation $r = r_{FP}$ is strictly valid with no selection and breaks down with very high h^2 and intense selection.

An *example* of prediction follows. Consider $N_m = 20$ males, $N_f = 100$ females, $n_f = 6$ individuals of each sex scored per full-sib family and $n_m = 30$ per half-sib family, random mating of selected parents and initial heritability 0.4. From standard statistical tables $i_m = 2.23$, $i_f = 1.50$, $x_m = 1.83$ and $x_f = 0.97$. Hence, $k_m = 0.89$, $k_f = 0.80$ and $k = (k_m + k_f)/2 = 0.84$. The heritability and intraclass correlation used in the prediction have to account for the "Bulmer effect". The heritability can be approximated by that after one generation of selection, $h_1^2 = h^2(1 - kh^2/2)/(1 - kh^4/2) = 0.36$. Under random mating, the asymptotic intraclass correlation of half sibs from parents of sex s can be approximated by $\rho_s^* \approx V_{bs}/(V_{bm} + V_{bf} + h^2/2 + [1 - h^2])$, where $V_{bs} \approx (h^2/4)(1 - k_s h_1^2)/(1 + k h_1^2)$, i.e. $\rho_m^* = 0.057$ and $\rho_f^* = 0.061$. Hence, the asymptotic intraclass correlation of families from parents of sex s is $\rho_s = \rho_m^* + \rho_f^*(N_s/N_f)$, i.e. $\rho_m = 0.069$ and $\rho_f = 0.118$. Therefore, $C_{mm}^2 \approx i_m^2 \rho_m = 0.34$, $C_{mf}^2 \approx i_f^2 \rho_m = 0.16$, $C_{fm}^2 \approx i_m^2 \rho_f = 0.59$, $C_{ff}^2 \approx i_f^2 \rho_f = 0.27$, $C_m^2 = [(C_{mm} + C_{mf})/2]^2 = 0.24$ and $C_f^2 = [(C_{fm} + C_{ff})/2]^2 = 0.41$. Now, $G_m = 1 - k_m h_1^2 = 0.68$, $G_f = 1 - k_f h_1^2 = 0.71$, $G = (G_m + G_f)/2 = 0.70$ and $Q = 2/(2 - G) = 1.54$. Finally, $\mu_{mm} = \mu_{ff} = 1$, $\mu_{mf} = 1/\mu_{fm} = 5$, $S_{mm}^2 \approx 1 - 1/n_m = 0.97$, $S_{mf}^2 \approx (N_f/N_m)(1 - 1/n_f) = 4.17$,

$S_{fm}^2 \approx (N_m/N_f)(1-1/n_m) = 0.19$, $S_{ff}^2 \approx 1-1/n_f = 0.83$, $S_{mm,mf} = S_{fm,ff} = 0$, $\alpha_m = -1/(2N_m-1) = -0.026$ and $\alpha_f = -1/(2N_f-1) = -0.005$. Substituting into the equation for N_{es} above, $N_{em} = 17.6$, $N_{ef} = 25.9$ and $N_e = 4N_{em}N_{ef}/(N_{em} + N_{ef}) = 41.8$. The value obtained is not exactly the same as in the table because predictions in the table for the HM scheme used a more accurate approximation of C_3^2 .

Compensatory mating increases effective size (see table above) by removing the cumulative effect due to selection. It does not affect, however, the variance of family size. By using heritability estimates in BLUP that exceed their true value, variance of family size can be reduced and effective size increased (Toro and Pérez-Enciso, 1990; Grundy *et al.*, 1994). This suggests that both procedures applied together could be very advantageous. The following table shows simulations of the two methods separately and in combination (Grundy *et al.*, 1994). Selection using an AM-BLUP was carried out in populations with N_m males and N_f females, where each male was mated to N_f/N_m females and n individuals of each sex were scored per full-sib family and generation. Selection was practised using the true heritability (0.2) or a higher value (0.4) both under random mating (*R*) and/or with compensatory mating (*CM*). Under a hierarchical mating scheme, compensatory mating is made by ordering individuals according to the sum of the total number of their selected half-sibs and the total number of their selected full-sibs with males given a weight N_f/N_m that of females to reflect their contributions.

N_m	N_f	n	<i>R</i> (0.4)			<i>CM</i> (0.2)			<i>CM</i> (0.4)		
			G_5	G_{10}	ΔF	G_5	G_{10}	ΔF	G_5	G_{10}	ΔF
10	10	3	-0	-0	-24	-0	-1	-15	-1	-0	-33
20	20	3	+1	-1	-23	-0	+0	-16	-1	-1	-35
20	20	9	-1	+3	-33	-0	+3	-15	-3	+3	-47
20	60	3	-1	-1	-26	+5	+0	-10	-4	-3	-37
20	120	3	-2	-1	-30	+1	+0	-12	-2	-2	-41

The table shows average responses to selection (G_t) in generation t and average rates of inbreeding (ΔF) between generations 5 and 10, as a proportional deviation from the corresponding value with random mating and using the true heritability. By using 0.4 in the AM-BLUP, rates of inbreeding are reduced by about 25%, for the range of cases simulated. Compensatory mating alone reduces rates of inbreeding by about 15%, on average. Finally, when both procedures are applied, rates of inbreeding are reduced by about 40%, showing that the two methods have approximately additive effects on the rate of inbreeding. Responses to generation 5 or 10 were very little reduced, if any, by either or both procedures.

Acknowledgements: We thank the AFRC and EEC for financial support and W. G. Hill for comments.

REFERENCES

- BULMER, M. G. (1971) *Am. Nat.*, 105: 201-221.
 CABALLERO, A. (1994) *Heredity* (in press).
 GRUNDY, B., CABALLERO, A., SANTIAGO, E. and HILL, W. G. (1994) *Anim. Prod.* (submitted)
 ROBERTSON, A. (1961) *Genet. Res.*, 2: 189-194.
 SANTIAGO, E. and CABALLERO, A. (1994) *Genetics* (submitted).
 TORO, M. A. and PEREZ-ENCISO, M. (1990) *Genet. Sel. Evol.*, 22: 93-107.
 WOOLLIAMS, J. A., WRAY, N. R. and THOMPSON, R. (1994) *Genet. Res.*, 62: 231-242.
 WRAY, N. R. and THOMPSON, R. (1990) *Genet. Res.*, 55: 41-54.
 WRAY, N. R., WOOLLIAMS, J. A. and THOMPSON, R. (1994) *Theor. Appl. Genet.* (in press).
 WRIGHT, S. (1931) *Genetics*, 16: 97-159.