Comparison of family and sib selection during selfing

E. M. TOMS, M. A. CORNISH & M. J. KEARSEY*

School of Biological Sciences, University of Birmingham, Birmingham B15 2TT, UK

Theory is presented and subsequently validated by computer simulation for predicting the difference in response to family and sib selection during selfing. The response is measured by the change in mean and variance of the distribution of inbred lines due to selection. A combined measure of the response is also given by the proportion of lines scoring higher or lower than a given target. Simulation of a highly heritable character under simple genetic control demonstrated the reliability of the theory. However, with a character of moderate heritability involving additive and nonadditive effects, and experimental sizes more common in practice, the differences between family and sib selection proved to be small. This was further confirmed by the sampling variation of the responses.

Keywords: family selection, heritability, prediction, selfing, sib selection, single seed descent.

Introduction

It is common for the breeder to practise family selection in the early generations of inbreeding. Prediction of the effect of such selection on the means of later generations has been described by a number of authors, such as Wricke & Weber (1986) and Snape & Simpson (1984). More recently, Cornish (1990a,b) examined the effects of selection during selfing on both the means and variances of later generations. The theory was developed for a simplified situation in which selfing by single seed descent (SSD) commences from the F_2 generation of a cross between two inbred parents. A single round of selection for a quantitative character is imposed at the N^{th} generation, either on the basis of individual scores at the F_2 or on family means in later generations. The selected lines are selfed by SSD to the F_∞ generation, at which stage the lines are assessed. The response is measured by the difference in mean and variance of the distribution of selected lines compared with the distribution of lines derived without any selection. The predicted response shown by the mean is:

$$R = ih_{F_N}^2 \sigma_{P_{F_N}},$$

and the reduction in the variance is given by:

$$\delta V = h_{F_N}^2 i(i-x)V_{A_{F_N}},$$

the selected generation, $h_{F_N}^2$ is the appropriate narrow

heritability, $\sigma_{P_{F_N}}$ is the phenotypic standard deviation, and $V_{A_{F_N}}$ is the additive genetic variance. Given that, in the absence of selection, the expected mean of the distribution of F_∞ lines is equal to the midparental value (m) and the expected variance is equal to the additive genetic variance (D) using the notation of Mather & Jinks (1982), then, following a single round of selection, the distribution of the selected lines is predicted by:

$$\begin{aligned} \text{mean} &= m \pm ih_{F_N}^2 \sigma_{P_{F_N}} \\ \text{variance} &= D - h_{F_N}^2 i(i-x)V_{A_{F_N}}. \end{aligned}$$

The appropriate estimates of these parameters depend on the type of selection applied.

In true family selection, those individuals which contribute to the family mean are themselves selected and, by implication, this requires that the individuals being tested are inbred to continue the lines. This poses practical problems in many outbreeding species. Rather than selfing the experimental plants in the field, it is common practice to continue the lines from sibs of the selected families which can be selfed under greenhouse conditions. In this case, the selected individuals do not contribute to the estimate of their family mean. This variant of family selection is termed *sib* selection by Falconer (1989).

The narrow heritabilities appropriate for predicting the response to selection in these two distinct circumstances for a random mating population are different (Falconer, 1989). The equivalent approach for selection during inbreeding has not been described, but the principles are the same.

*Correspondence.

The general form of the narrow heritability of family means in the N^{th} generation is defined as:

$$h_{F_N}^2 = V_A/V_P,$$

where, in the notation of Mather & Jinks (1982),

$$V_A = \left[1 - \left(\frac{1}{2} \right)^{N-2} \right] D + \left(\frac{1}{n} \right) \left[\left(\frac{1}{2} \right)^{N-1} \right] D \text{ and}$$

$$V_P = \left[1 - \left(\frac{1}{2} \right)^{N-2} \right] D + \left[\left(\frac{1}{2} \right)^N - \left(\frac{1}{2} \right)^{2N-2} \right] H + \left(\frac{1}{n} \right) \left[\left(\frac{1}{2} \right)^{N-1} D + \left(\frac{1}{2} \right)^N H + E \right],$$

D being the additive genetic variance, H the dominance variance, E the environmental variance and n the family size. In the special case of the F_2 , this simplifies to

$$h_{F_2}^2 = \frac{\frac{1}{2}D}{\frac{1}{2}D + \frac{1}{4}H + E}.$$

The numerator is the covariance between F_N family means and the F_∞ generation, and contains both rank 1 and rank 2 components (Cornish, 1990a). It is composed of the additive genetic variance between F_N family means (rank 1) as well as a fraction of the additive genetic variability between individuals within families (rank 2). The magnitude of this fraction is dependent on the family size: the larger the family, the smaller the fraction of this within-family variance. This heritability is appropriate for predicting the response to true family selection. If, however, the selected individuals do not contribute to the estimate of their family mean and variance, which is the case for sib selection, then it is not appropriate to include the within-family additive genetic variation which is specific to the family members actually assessed. Thus the narrow heritability of F_N family means applicable for predicting the response to sib selection is given by:

$$h_{F_N}^2 = V_A/V_P,$$

where

$$V_A = \left[1 - \left(\frac{1}{2} \right)^{N-2} \right] D$$

and V_P is as above.

The validity and accuracy of these formulae to describe the difference in response to family and sib selection was assessed by computer simulation. The difference was examined in terms of the mean and variance of the distribution of lines. As a combined measure of the response, the proportion of lines scoring higher or lower than a target was also assessed, by extending the method given by Jinks & Pooni (1976) to selection responses. The variability of individual responses was also examined. A method of calculating the confidence interval of the estimated response is given in detail by Toms (1992).

Method

A programme of selfing with selection was simulated using the computer program written by Cornish (1990a,b) for modelling plant breeding situations. Two inbred parents were crossed to produce an F_1 hybrid which was then selfed to generate a set of F_2 individuals. SSD was initiated from this F_2 to generate a set of F_6 lines, each line descending from a different F_2 individual. Selection for low score was imposed on the simulated quantitative character at either the F_3 or F_4 generation, on the basis of family means which were obtained from a fully randomized trial. The effects of selection were assessed at the F_6 generation. Independent runs of the simulation were performed for a range of selection intensities from 1 per cent to 100 per cent, and each simulation was repeated 200 times.

The simulations were carried out for two different model situations. The first tested the accuracy of the theory under hypothetical conditions which maximized the potential differences between family and sib selection. The second examined the theory under experimentally more realistic conditions.

Thus the first set of simulations used a character with a narrow heritability of 100 per cent, controlled by 20 unlinked loci of equal additive effects, with no dominance. The second set used a more realistic character with a moderate heritability of 30 per cent, controlled by 10 unlinked genes each contributing equal additive and dominance effects, with dominance acting for reduced score. There were no epistatic effects in either character.

For the first situation, the simulated selection programme was very simple. A sample of 1000 lines was generated, and each family was represented by just one sib in the selection trials, to maximize the potential difference between family and sib selection. The selected lines were assessed at the F_6 generation; again each line was represented by a single individual.

For the second situation, a smaller sample of 100 lines was generated, and each family was represented

by three sibs in the F_3 trial or five sibs in the F_4 trial. A sample of the F_1 generation was also included to provide an estimate of the environmental variance. Assessment at the F_6 generation used a family size of five. The data from the F_3 and F_4 trials were also analysed using standard biometrical methods (Mather & Jinks, 1982) to estimate the parameters required to predict the selection response. Thus, both the *exact* predictions determined from the parameters defining the character and the *estimated* predictions derived by analysis from the simulated trial data were compared with the observed selection responses. The response

was also assessed as the proportion of lines scoring less than the F_1 mean. The variability of individual responses was examined in terms of the 95 per cent confidence interval.

Predicted responses were calculated using the formulae presented previously. Values of i and x were obtained from Harter (1961) when selection was from populations of 100 or less, and from tables given by Falconer (1989) for populations larger than 100. Because of the precision of highly repeated simulations, the F_∞ variance (D) overestimates the variance of the F_6 generation, so the exact expectations of the

Fig. 1 Comparison of the expected and observed means and variances of F_6 lines derived from the F_3 or F_4 generation following family or sib selection for the simulated character of 100 per cent heritability. The expected values are indicated by curves — the means by continuous lines and the variances by broken lines — and the observed values are shown by points.

variances were used: for the first set of simulations the expected variance between F_6 families was $\frac{31}{32}D$ (with a family size of one); for the second set the between-family variance was $\frac{15}{16}D$. The observed selection response was measured by the difference in mean and between-family variation of the selected and unselected lines, averaged over the 200 repeat runs for each selection intensity.

Results and Discussion

The expected and observed responses to family selection and sib selection at the F_3 or F_4 generation are compared in Fig. 1 for the character of 100 per cent heritability. The results clearly and consistently demonstrated that true family selection induced a greater response than sib selection, in both the mean and the variance, as expected from the theory. This difference, however, was narrower the later the generation of selection. Comparison of the observed and expected responses showed that, without exception, the observed means coincided almost exactly with the expected means. The observed variances were also very close to the expected values at the lower selection intensities, but there was a slight departure from expectation under very intense selection. This was most probably because of the finite number of loci controlling the character. The distribution of a character is only truly normal if it is controlled by an infinite number of genes, while a character controlled by a finite number of loci has a binomial distribution. In practice, however, for a character controlled by a large number of loci, the binomial distribution closely approximates the normal distribution except at the tails. It is thus under very intense selection that the non-normality of the distribution is apparent, and hence the departure from expected at these intensities.

The results of the more realistic simulations with 30 per cent heritability were not so clear-cut as those of the more theoretical simulations. The exact and estimated predictions together with the observed responses to F_3 or F_4 selection are displayed in Fig. 2. It is clear from the expected curves that the differences between family and sib selection were substantially narrowed when heritability is moderate compared with the more highly heritable character presented previously. Nevertheless, the effect of increased family size was also confounded with this effect.

Predictions of the mean and variance obtained from the simulated trial proved to be quite accurate. The graphs do indicate, however, that there may be problems in distinguishing between family and sib selection particularly at later generations of selection and with larger family sizes. Nevertheless, the results confirm

the accuracy of the estimation procedures used for predicting the response to selection.

The observed responses in the mean and variance also fitted the expected values quite closely. Some divergence in the variances was apparent at higher selection intensities, which may be attributed to the small number of loci controlling the character (Cornish, 1990a).

As a combined measure of the response by the mean and variance, the proportion of lines scoring lower than the F_1 mean was calculated. Figure 3 shows a fairly close correspondence between the exact, estimated and observed proportions, although the responses to family and sib selection were not unambiguously identifiable. It appears that the divergent variances noted previously in Fig. 2 had little effect on this combined measure of the response.

Of greater practical consequence than the average responses presented thus far is the variability of individual responses. The 95 per cent confidence intervals of the estimates and the intervals including 95 per cent of repeat observations following sib selection are also shown in Fig. 3. The plots illustrate the extent of the variability of individual runs; both the estimated and observed proportions varied widely around the exact expectations, and this variability increased quite dramatically as the number of lines was reduced under increasingly intense selection. The predicted confidence intervals are very close to those observed under mild selection but increasingly underestimate the variability under more intense selection. Given these large confidence intervals, the differences between family and sib selection noted from the average responses were clearly not significant.

Conclusions

The difference between true family selection and sib selection was shown to be accurately described by the theory presented, validating use of the appropriate estimate of the narrow heritability in predictions. The magnitude of the difference is, however, very much dependent on the heritability of the character, the generation of selection and on the family size in the selection trial, with the difference becoming narrower the lower the heritability, the later the generation and the larger the family. In reality the difference is more likely to be of theoretical than practical importance, given the wide limits to the sampling variation which were apparent.

In addition to these conclusions, this study confirmed the accuracy of the response equations presented by Cornish (1990a) for predicting the effects of

selection during selfing on both the mean and the variance of the distribution of derived lines. This theory is crucial for designing optimal breeding programmes (Cornish, 1990b). It was further demonstrated that the selection response can be conveniently described by the proportion of lines scoring higher or lower than a specified target, by extending the method given by Jinks & Pooni (1976). This may find particular use in comparing predicted and observed responses

estimated in different environments, relative to a target scored in each environment. The method given by Toms (1992) for estimating the variability of this response was also shown to be quite accurate and is vital for statistical comparison of predicted and observed responses.

Clearly, the absence of complicating factors in the design of these simulations greatly enhanced the clarity of the results. In reality, of course, various other effects

Fig. 2 Comparison of the exact, estimated and observed means and variances of F_6 lines derived from the F_3 or F_4 generation following family or sib selection for the simulated character of 30 per cent heritability. The expected values are indicated by curves — the means by continuous lines and the variances by broken lines — and the observed values are shown by points.

Fig. 3 Comparison of the exact, estimated and observed proportions of F_6 lines scoring less than the F_1 mean, following family or sib selection at the F_3 or F_4 generation for the simulated character of 30 per cent heritability. The expected values are indicated by the continuous lines, the observed values are shown by the points. For sib selection, the intervals including 95 per cent of the estimated and observed proportions are also shown.

would also be present, such as non-allelic interactions, linkage, genes of unequal effect and the presence of $G \times E$ interaction. Nevertheless, Cornish (1990a,b) found that biases resulting from epistasis were still smaller than the sampling errors.

References

- CORNISH, M. A. 1990a. Selection during a selfing programme. I. The effects of a single round of selection. *Heredity*, **65**, 201–211.
- CORNISH, M. A. 1990b. Selection during a selfing programme. II. The effects of two or more rounds of selection. *Heredity*, **65**, 213–220.
- FALCONER, D. S. 1989. *Introduction to Quantitative Genetics*, 3rd edn. Longman Scientific & Technical, Harlow.
- HARTER, H. L. 1961. Expected values of normal order statistics. *Biometrika*, **48**, 151–165.
- JINKS, J. L. AND POONI, H. S. 1976. Predicting the properties of recombinant inbred lines derived by single seed descent. *Heredity*, **36**, 253–266.

- MATHER, K. AND JINKS, J. L. 1982. *Biometrical Genetics*, 3rd edn. Chapman & Hall, London.
- SNAPÉ, J. W. AND SIMPSON, E. 1984. Early generation selection and rapid generation advancement methods in autogamous crops. *Efficiency in Plant Breeding. Proceedings of the 10th Congress of EUCARPIA*, pp. 82-86. Wageningen, Netherlands.
- TOMS, E. M. 1992. *An Investigation into the Effects of Selection within an Inbreeding Programme in Sunflower*. Ph.D. Thesis, University of Birmingham.
- WRICKE, G. AND WEBER, W. E. 1986. *Quantitative Genetics and Selection in Plant Breeding*. Walter de Gruyter, Berlin.