THE PRODUCTION OF SUPERIOR LINES AND SECOND CYCLE HYBRIDS BY INBREEDING AND SELECTION

Keywords: Inbreeding; Selection; Efficiency of improvement programmes; Line and cross prediction

H A POONI, M A CORNISH, M J KEARSEY
and M J LAWRENCE*

The realization that heterosis is due chiefly to the dispersion of dominant genes between the parents of a cross rather than to the overdominance of these genes has led to the important practical conclusions that, in principle, heterosis should be fixable in inbred lines. It follows, therefore, that there is no longer any genetical justification for regarding the improvement of outbreeding species any differently from that of inbreeding ones, so that all crops and stock can be improved by inbreeding programmes. In the past, breeders have had to employ an essentially empirical approach in carrying out improvement programmes which involve inbreeding. Recent work, however, has led to an understanding of the theoretical basis of the advances to be expected under inbreeding, both with and without selection, so that it is now possible for the breeder to design and carry out improvement programmes in a systematic way and, hence, to predict, from the early generations of a pedigree, the likelihood of obtaining recombinant inbred lines or hybrids that have the desired level of performance. In this paper we review this theory and discuss its application to the practical task of crop improvement.

---

* School of Biological Sciences University of Birmingham B 15 2 TT, U K

Paper delivered by Dr M J Lawrence on 27 January 1989 as a PORIM seminar during his visit to the institute (23-29 January 1989)

INTRODUCTION

This is the third time during the last decade that we have had the pleasure of addressing an audience in Malaysia on the application of biometrical genetics to plant and animal breeding. In our previous papers (Jinks and Lawrence, 1979, 1983) we drew attention to the fact that virtually all biometrical geneticists now agree that the chief cause of heterosis is the dispersion of genes in the parents of hybrids that display directional dominance, with or without directional epistasis or non-allelic interaction, in the increasing direction, rather than, as has been supposed in the past, to overdominance of these genes (Jinks, 1981, 1983). This conclusion has important implications for all breeders, for it suggests that the performance of the best homozygote should at least equal and, unless dominance is complete, exceed that of the best heterozygote that can be extracted from a cross; and that heterosis should be fixable in inbred lines. It follows, therefore, that there is no longer any genetical justification of treating the improvement of outbreeding species differently from that of inbreeding ones; it also follows that, with the constraints imposed by the biology of the organism in question, all may and, indeed, can be improved by inbreeding programmes, a conclusion which brings a very welcome unity to the practice of both plant and animal breeding.

Now the chief breeding programme used to improve those crop plant species that are predominantly self-pollinating is, of course, the pedigree method, in which two or more inbred parents are crossed to produce a segregating F_2 populations from which recombinant inbred lines are extracted by successive generations of self-pollination. Though the way in which the pedigree method is applied varies in detail from one crop to another, most breeders attempt to carry out selection during the course of inbreeding, which usually involves the visual selection of individuals in the early generations, and of families (*e.g.* ear rows) in the more advanced generations of the pedigree. Successful breeders initiate many pedigrees by making many crosses between parental inbreds. In practice, however, the great majority of these pedigrees are discarded for one reason or another, some as early as the F_1 stage and others later; that is, very few pedigrees run full term to produce new and improved varieties. Hence, whilst the use of the pedigree method with self-pollinating crops has resulted in the introduction of hundreds of new varieties, from the genetical point of view its efficiency cannot be regarded as wholly satisfactory on a per pedigree basis. We may ask, therefore, what can be done to improve this efficiency?

In considering this question, it is convenient to recognize three key stages in any pedigree. Thus, success of any pedigree depends on the breeder:

(i) crossing the right parental inbred lines;

(ii) choosing the best crosses among the many made between parental inbreds; and

(iii) exploiting the genetical variation of the F_2 generation of the handful of promising crosses as efficiently as possible.

There is little that geneticists can say about the choice of parents, for this obviously depends on the judgement and experience of the breeder and, in particular, on his knowledge of the various ways in which current varieties of his crop fall short of what is desirable. Generally, however, the breeder will want to choose parents the attributes of which are complementary, the performance of which is satisfactory, and which are not too closely related, so as to increase the chance that they differ at a number of loci which control characters of interest. Genetical theory, however, can be used with advantage when considering how to raise efficiency in stages two and three of the pedigree method, for stage two involves the recognition of the genetic potential of a cross, while in stage three the question is how best to realize this potential in the form of one or more recombinant inbred lines that can be extracted from the cross.

We have, of course, discussed this question of efficiency in our previous addresses to this audience. We make no apology, however, for returning to the theme because, as a result of work carried out in Birmingham over the past twenty years or so, there is now a very extensive and detailed theory which can be used by any breeder who wishes to obtain an answer to this important, practical question. In the present paper we first consider the prediction (using information from the early generations of a pedigree) of the properties of the distribution of recombinant inbred lines that can be extracted by single seed descent or dihaploidy from a single cross between a pair of parental inbred lines, before moving on to the prediction of those extracted from three and four parent crosses. We shall then consider the prediction of the properties of the second cycle hybrids that can be obtained by crossing these recombinant inbred lines. Finally, we shall discuss the circumstances in which it is worthwhile practising some selection during the production of the inbred lines.

Our chief purpose in this paper is to show what can be done; that is, we shall be mainly concerned with principles. In order to leave no doubt, however, that these prediction techniques can be applied to the full range of genetical situations encountered by the breeder, we shall be giving a little more of the genetical detail underlying the theory than in our previous papers.

THE PROPERTIES OF RECOMBINANT INBRED LINES EXTRACTED FROM A SINGLE CROSS

We consider, first, the properties of the distribution of the recombinant inbred lines that have been produced by single seed descent from the F_2 generation of a two parent cross without selection. If the genes for which the parental inbreds differ display additive and dominance effects only, and the character in question is normally distributed, the distribution of the means of the recombinant inbred lines that can be extracted from the cross is also normal; in the notation of Mather and Jinks (1982), there is a mean m and variance D, where D is the additive genetic component of variance. Estimates of m and D can be obtained from a variety of sources, but the experiment which provides, in general, the most reliable estimates, and also allows tests of assumptions about the genetical architecture of the character, is one involving the basic generations of the cross (P_1, P_2, F_1, F_2, and the backcrosses, B_1 and B_2) and a set of triple test crosses, which are produced by backcrossing each of a random sample of, say, 25-35 F_2 individuals to P_1, P_2 and the F_1 (Kearsey and Jinks, 1968; Pooni and Jinks, 1976). The standard normal distribution theory may then be used to calculate the proportion of lines the mean performance of which equals or exceeds that of any desired standard, St. This proportion is found from tables of the one-tailed normal deviate, z, where

$$z = \frac{St - m}{D}$$

assuming, as will generally be the case, that the breeder is interested in selecting for greater expression of the character (Jinks and Pooni, 1976). The standard, St, may be an existing commercial variety, the parent of greater expression, P_1, or, where it displays heterosis, the F_1 of the cross.

In practice, the two most likely causes of departure from the simple additive dominance model of inheritance are epistasis, or non-allelic interaction, and linkage disequilibrium. Since epistasis can be easily detected in the basic generations of the cross and, indeed, can be accommodated by fitting an extended model to the means of these generations (Mather and Jinks, 1982), the breeder will be aware of its presence in his material. While, in the presence of epistasis, the mean of the distribution of inbred lines is still m, its variance is no longer D, but $D + I$, where I is the additive x additive component of genetic variance. It is not possible to obtain precise estimates of $D + I$ from the basic generations of the cross. The estimate of D from the triple test cross progenies, however, provides a closer approximation to the corresponding statistic for the inbred lines than those obtained from any other source (Pooni and Jinks, 1979; Pooni, Jinks and Cornish, 1977). Strong directional epistasis can also skew the distribution of inbred line means but, in practice, this is unlikely to give a significant departure from normality unless the character is also highly heritable. Hence, the proportion of recombinant inbred lines that equal or exceed any particular standard can, in the presence of epistasis, be calculated in the same way as before, except that the estimate of *m* must be obtained by fitting an extended model to the means of the basic generations. Turning now to the second most likely cause of departure from the simple additive-dominance model, linkage, as such, has no effect on either the mean of the distribution of recombinant inbred lines, nor on its variance. Linkage in the guise of linkage disequilibrium, however, does affect the magnitude of the additive genetic variance, *D* (Jinks and Pooni, 1976, 1982). Thus, estimates of *D* from the early generations of a pedigree (say the F_2) are expected to be larger than those obtained from the more advanced generations (say F_6) when the linkage disequilibrium is predominantly in the coupling phase, and smaller than the latter when this disequilibrium is predominantly in the repulsion phase. This is because there are more opportunities for recombination with each generation of the pedigree.

The breeder is more likely to encounter linkage disequilibrium in the repulsion phase. Thus, in practice, he will often choose, as parents of a cross, lines the attribute of which are complementary, with a view to combining all the desirable attributes in one recombinant inbred line that is to be extracted from the cross. In other words, the genes with desirable expression for each character are likely to be dispersed over the two parents, rather than being associated in one of them. In so far as some of these dispersed genes are linked in their inheritance, they will be linked in the repulsion phase. With only one round of recombination between the parental generation and the F_2, these genes will tend to be still linked in the repulsion phase in the latter generation. In short, the breeder's choice of parents will often tend to generate linkage disequilibrium in the repulsion phase in the F_2 generation of the cross. Nevertheless, the presence, magnitude and phase of linkage disequilibrium can be determined in the early generations of the pedigree (Perkins and Jinks, 1970). So, as with epistasis, the breeder will be aware, from his data, of the presence of this disturbance. The prediction of the proportion of recombinant inbred lines of which the performance equals or exceeds any particular standard can be made in the same way as previously described, provided that it is borne in mind that this prediction will slightly underestimate the probability of success in obtaining a superior line when linkage disequilibrium is predominantly in the repulsion phase.

As might be expected, linkage disequilibrium between pairs of loci displaying epistasis affects both the mean and variance of the distribution of recombinant inbred lines. Again, this situation can be detected in the early generations of a pedigree by comparing the means of the B_1, B_2 and F_2 families with the L_1, L_2 and L_3 families of the triple test cross (Jinks, 1978). The information, from these comparisons can indicate the likely deviation of the mean of the recombinant inbreds from *m*; however, the contribution which linkage disequilibrium and epistasis make to the mean and variance of their distribution is rather complex, so that, in practice, it is not possible to obtain precise estimates of the magnitude of either of these effects from the early generations of the pedigree (Jinks and Pooni, 1981). The best that can be done in these circumstances is to obtain estimates of these parameters from the same source as before, namely, *m* from the basic generations, and *D* from the triple test cross progenies. The proportion of recombinant inbred lines which equal or exceed any desired standard can then be calculated as before, though predictions in these circumstances are naturally expected to be somewhat less precise than those made for characters the genetical architecture of which is simpler.

The effect of epistasis and/or linkage disequilibrium on the prediction of the properties of the recombinant inbred lines that can be extracted from a cross has been discussed in some detail, not only because both are encountered in practice, but also because we wish to make it clear that the application of this useful procedure is not limited to characters which display only additive and dominance effects.

There is one further complication about

which the breeder needs to be aware. The estimates of the statistics used to predict the properties of the recombinant inbred lines are obtained from experiments carried out at the F_2 or F_3 stage of the pedigree, whereas the actual performance of the lines will be assessed in a trial carried out at the F_6 stage; that is, these experiments are necessarily carried out in different seasons and, hence, in different environments. Provided the effects of genotype and the environment are independent, this causes no problem, for the breeder can measure the effect of the environment by including the parents, P_1 and P_2, in the trial of the F_6 lines. If, however, the character of interest displays genotype x environment interaction, because different genotypes respond differentially to a given change of environment, prediction will no longer be very precise; however, only rather rarely will the relative magnitude of genotype x environment interaction be so great as to make predictions worthless. This problem is not, of course, peculiar to the present procedure for, in the presence of genotype x environment interaction, comparisons between experiments become less straight forward than in its absence, though the magnitude of the problem depends on the nature of the interaction: see Jinks and Lawrence (1979) for a discussion of this point.

DIHAPLOIDS

The production of recombinant inbred lines by single seed descent can be accomplished with relative ease in a small-grain cereal crop like rice, in which three generations per annum can be raised without much difficulty. In these circumstances, it is possible, if the performance of the recombinant inbred lines is determined in an appropriately designed trial at F_6, to complete the task in two years. In crop like oil palm, on the other hand, the same task would take very much longer, even if it were possible to carry out the single seed descent phase of the programme in the nursery and the generation time was reduced to two years. With crops like this, it would clearly be advantageous to produce the homozygous lines by dihaploidy, rather than by single seed descent, assuming, of course, that this alternative was available.

In the absence of linkage disequilibrium and selection, conscious or otherwise, lines obtained by dihaploidy should be identical with those produced by single seed descent. However, the haploids, from which the dihaploids are derived, are usually produced from anthers taken from the F_1 plants of a cross, so that there will be only one round of recombination between the parental generation and the inbred products, whereas, there will be several such rounds when the latter are obtained by single seed descent. It follows, therefore, that, when linkage disequilibrium is present, the products of dihaploidy and single seed descent are not expected to be the same: the latter display a greater range of variation than the former when the disequilibrium is predominantly in the repulsion phase (Jinks and Pooni, 1981a). Hence, one is less likely to obtain useful transgressive variation with lines derived by dihaploidy than those obtained by single seed descent. This problem can be considerably alleviated, however, by deriving the haploids from plants of the F_2 or F_3 generation, in which case differences between the products of dihaploidy and single seed descent become smaller, the former now being the offspring of two or three rounds of recombination, rather than one. Thus, where anther culture has been fully developed with a crop, dihaploidy can be a very useful addition to the array of techniques available to the breeder.

THREE AND FOUR PARENT CROSSES

Hitherto, our discussion has been confined to the case where a pedigree is initiated by a single cross betwen two parental inbred lines. In practice however, the breeder may be unable to find a pair of lines which, between them, possess all the attributes required in a new and improved variety; or he may feel that the F_2 from a single cross is unlikely to contain sufficient genetical variation to allow him a reasonable chance of success in meeting the objectives of the breeding programme. In these circumstances, the breeder will want to initiate pedigrees from three or four parent crosses, so that the foundation cross is of the type $(P_1 \times P_2) \times P_3$ or $(P_1 \times P_2) \times (P_3 \times P_4)$, rather than $P_1 \times P_2$, when only two parents are involved.

The genetics of three and four parent crosses is considerably more complex than that of two parent crosses: in these only two alleles can be segregating among the descendants, whereas three may be segregating among those of three parent crosses, and four among those of four parent crosses. Despite this complexity, however, it is still possible to predict the parameters of the distribution of the recombinant inbred lines that can be extracted from such crosses using estimates obtained from the basic generations of the related single crosses; or from the first two generations obtained by self-pollinating a random sample of individuals of a four parent cross (Pooni and Jinks,1985b; Pooni, Jinks and Yohannes, 1985;1986).

PREDICTIONS INVOLVING TWO OR MORE CHARACTERS SIMULTANEOUSLY

The breeder will nearly always have to take into account several characters, rather than one, when attempting to improve a crop. The procedures we have discussed so far will lead to useful joint predictions for two or more characters simultaneously only if these characters are independent. In practice, this will rarely be the case. We need, therefore, to be able to predict the probability of obtaining a recombinant inbred line from a cross that is superior for all of the characters of interest to the breeder, allowing for the fact that many, if not all, of these characters will not be independent. The procedure described in the section entitled 'The Properties of Recombinant Inbred Lines Extracted from a Single Cross' above can be easily extended to handle two or more characters simultaneously, provided that estimates of the additive genetic correlation between pairs of these characters are available (Pooni and Jinks, 1978). Estimates of these correlations can be obtained from the triple test cross progenies.

SECOND CYCLE HYBRIDS

We have so far assumed that the end-product of a breeding programme is an inbred line the performance of which is, in one way or another, superior to that of any existing variety. For reasons given earlier, there are no genetical grounds for doubting that this is possible, even if the best existing variety is an F_1 hybrid; however, in practice it may turn out difficult to obtain a superior inbred line from just one cycle of inbreeding.

There are two chief reasons why the breeder may find himself in this position. First, as explained earlier, it is likely that some of the genes with high expression for the characters of interest are dispersed over the parents of the cross and that some of these dispersed genes are linked in their inheritance: in this case, they will be linked in the repulsion phase. If, furthermore, this linkage is quite tight, it may be difficult to produce a sufficiently large number of lines by single seed descent from a cross to ensure that, at a reasonable probability, at least one of these lines is recombinant for these pairs of tightly linked genes. Indeed, it is worth pointing out that pairs of genes tightly linked in the repulsion phase are arguably more frequent in crops that have been greatly improved in the past than in those which have received less attention. This is because genes which assort freely or are only loosely linked will be the first and those which are tightly linked in the repulsion phase will be the last to be fixed during the course of selection.

The second reason why it may be difficult to obtain an inbred line that has the desired level of performance after only one cycle of inbreeding is that, while the breeder may be able to recover lines that excel for each of the characters of interest, none of them may be satisfactory when these characters are considered jointly. For independent characters, the probability of obtaining a line of the desired multiple or aggregate phenotype is equal to the product of the probabilities of obtaining lines that are superior for each character of the aggregate phenotype considered in turn. Unless these individual probabilities are quite large, their product will be small; and for characters that are negatively correlated, which are not infrequently encountered in practice, the probability of obtaining a line with the desired aggregate phenotype is even smaller. This is, of course, exactly the same problem that the breeder has to face when using the method of independent culling levels in a population of an outbreeding organism (Simmonds, 1979).

Now the obvious solution to both of these problems is to cross the best lines produced by the first cycle of inbreeding, thereby producing superior second cycle hybrids from which a second cycle of recombinant inbred lines can be extracted. Provided that the parents of this second cycle of inbreeding are chosen with care, so that they differ in at least some genes of interest, the probability of obtaining the desired combination of genes and/or characters in at least one line should be greater than in the first cycle; and, of course, this probability can be calculated on information from the early generations of the second cycle pedigrees, just as it can for the first cycle. This is, however, a long-term solution to the problem, and the breeder may well have to justify his activities in the short term. The short-term solution is to use the best of the second cycle hybrids as intermediate end-products of the breeding programme, until such time as superior inbred lines are available.

Just as it is possible to predict the properties of the recombinant inbred lines that can be extracted from a cross on information obtained from the early generations of the pedigree, so it is also possible to predict those of the second cycle hybrids produced by crossing these lines (Toledo, Pooni and Jinks, 1984a,b; Pooni, Jinks and Yohannes, 1985). In the absence of selection, the array of gametes produced by these inbred lines should be identical with the array of gametes produced by the first cycle F_1. It follows, therefore, that the population of second cycle hybrids produced by crossing the inbred lines in all possible ways should have the same mean and variance as the F_2 generation of the pedigree; that is, assuming that the genes controlling the character in question display additive and dominance effects only, the mean of the distribution of second cycle hybrids is expected to be $m + {}^1/_2\,[h]$ and its variance, when each hybrid combination is represented by just one plant, ${}^1/_2D + {}^1/_4H + E$, where $[h]$ is the dominance component of generation means and H and E are, respectively, the dominance and environmental components of phenotypic variance (Mather and Jinks, 1982). If, as will usually be the case, each hybrid combination is represented by a plot of plants, rather than just one, the environmental component can be ignored and the expected variance of the distribution of second cycle hybrids becomes ${}^1/_2D + {}^1/_4H$. The standard normal distribution theory can then be used as before to calculate the proportion of second cycle hybrids the performance of which equals or exceeds any particular standard from the equation:

$$z = \frac{St - (m + {}^1/_2[h])^{1/2}}{({}^1/_2D - {}^1/_4H)}$$

In the presence of epistasis, the mean and variance of the distribution of the means of second cycle hybrids are no longer $m + {}^1/_2[h]$ and ${}^1/_2D + {}^1/_4H$, respectively. However, because the genetical composition of the population of second cycle hybrids is, in the absence of linkage disequilibrium and selection, identical with that of the F_2 generation from which they were derived, the mean and variance of the latter can be used to predict the corresponding parameters of the former (Toledo, Pooni and Jinks, 1984 a,b).

Linkage disequilibrium has a similar effect on the variance of the distribution of second cycle hybrids as on that of the distribution of the inbred lines from which they are obtained: that is, if such disequilibrium is predominantly in the repulsion phase in the cross giving rise to the pedigree, the genetic variance of the F_2 generation is expected to be smaller than that of the second cycle hybrids, so that, in using the equation above, the breeder will tend to underestimate the proportion of hybrids that equal or exceed any particular standard (Toledo, Pooni and Jinks, 1984 a,b.)

In the presence of both epistasis and linkage disequilibrium, the mean and variance of the distribution of second cycle hybrids become quite complex and can no longer be very easily estimated by statistics obtained from the early generations of the pedigree. But because, as previously mentioned, the method of prediction is based on standard normal distribution theory, it is quite robust, so that the breeder is unlikely to be seriously misled, even when both of these disturbances are present in his data.

Our discussion on the uses of second cycle hybrids has so far been confined to the first of the problems mentioned at the beginning of this section, namely, the difficulty of obtaining inbred lines which are recombinant for genes linked tightly in the repulsion phase. Crossing a pair of inbred lines in which tightly linked genes are dispersed is one obvious and direct way of collecting them in one hybrid genotype. If, furthermore, these genes display dominance in the increasing direction, we expect that their F_1 will display heterosis for the character in question. Indeed, it is worth pointing out that, until pairs of such genes recombine, they will behave as a single segregational unit or effective factor (Mather and Jinks, 1982) which, even if dominance at each locus is less than complete, will apparently display overdominance. However, since in these circumstances, the apparent overdominance is a property of two loci, rather than one, it is better described as pseudo-overdominance (Simmonds, 1979).

A second point worth making about second cycle hybrids is that they can be produced by crossing the inbred lines to the F_1 of the cross founding the pedigree, as an alternative to crossing them among themselves in diallele fashion. While this alternative is a more economical way of producing these hybrids, it suffers from the disadvantage that they cannot be replicated, as they can when produced by crossing inbred lines, because the F_1 is, of course, heterozygous (Pooni, Jinks and Toledo, 1985). With some crops, however, it may be possible to replicate genotypes by vegetative propagation.

The second problem mentioned in the earlier part of this section concerned the difficulty of obtaining a satisfactory level of performance for each of the characters that contribute to the aggregate phenotype in one inbred line. The obvious way of tackling this problem is to cross two or more parents that, between them, exhibit the desired level of expression for each of the component traits. Having done this, however, the breeder will wish to ascertain the probability of obtaining a second cycle hybrid of the required level of performance with respect to the desired aggregate phenotype; this can be calculated in essentially the same way for hybrids as for inbred lines.

SELECTION DURING THE COURSE OF INBREEDING

Hitherto, we have confined discussion to circumstances where inbreeding is carried out without selection, conscious or otherwise. The advantage of the single seed descent procedure is that, because only a small amount of seed is required from each plant in each generation during the inbreeding phase of the pedigree and no attempt is made to carry out selection, plants can be raised at a high density in conditions in which the sole purpose is to ensure that the interval between successive generations is as short as possible. However, once the material has been advanced to the F_5 generation, one or two plants of each line can be raised at a lower density and encouraged to set sufficient seed for a trial at agricultural density in the following generation. Provided that each line is adequately and appropriately replicated in this trial, differences between their means should be largely genetic, so that the breeder has no difficulty in identifying those with the desired level of performance.

Many breeders, however, carry out selection during the inbreeding phase of the pedigree for two related reasons. First, it is argued that to defer selection until the advanced generations of a pedigree results in the retention of much material that is of little commercial value, and that the limited resources available to the breeder would be better employed in raising material that shows, in his opinion, some promise. Second, many breeders feel that it is possible to carry out effective selection during the course of inbreeding. In the early generations of the pedigree, when segregation of the genes for which the parental lines differ is at its greatest within families, individuals of superior phenotype are often selected by eye; later, when most of the genetical variation occurs between, rather than within, families, selection, again by eye, involves families (*e.g.* ear-rows) rather than individuals. Direct selection for yield, however, is usually deferred until F_4 or F_5 when enough seed can be produced for a small-scale trial at agricultural density.

Now the accuracy with which the breeder can identify plants of superior genotype by choosing those of superior phenotype depends, of course, on the heritability of the character for which selection is being practised. Where the parents of a cross differ for genes that have a major effect on the phenotype, such as those determining dwarfism or single gene resistance in some cereal crops, there is little doubt that the breeder can identify individuals of superior genotype from their phenotype, because the variation produced by the segregation of such genes in the F_2 generation of the cross is discontinuous; in other words, characters of this kind have heritabilities which are equal or close to unity.

The great majority of characters for which the breeder practices selection are, of course, not of this kind, because they are determinied by many, rather than one or two genes, the individual effect of which on the phenotype is small, and modifiable by the environment. The variation produced by the segregation of these genes in an F_2 family is continuous, rather than discontinuous, and the breeder can no longer accurately identify individuals of superior genotype from their phenotypes, because the heritability of these quantitative characters is typically much less than one. We may ask, therefore, whether there is any genetical justification for attempting to carry out selection for such characters during the inbreeding phase of a pedigree?

This question about the efficiency of selection in the early generations of a pedigree has been investigated in some detail in two very diverse crosses of *Nicotiana rustica*. While it is of little commercial value, more is probably known about the inheritance of quantitative characters in this species than in any other, because it has been used to test ideas and procedures in biometrical genetics for nearly 50 years. The empirical evidence from experiments carried out with this species suggests that, for characters of low to intermediate heritability, the value of selection in the early generations of a pedigree is low (Brumpton, Boughey and Jinks, 1977; Jinks, Jayasekara and Boughey, 1977; Boughey and Jinks, 1978; Jinks and Pooni, 1981b). More recently, Pooni and Jinks (1985a), have examined the value of applying a single cycle of selection in the F_2, F_3, F_4 or F_7 generation of a pedigree for each of the seven characters the narrow-sense heritabilities of which ranged from 0.34 (leaf width) to 0.69 (flowering time). They found that, although selection was effective in every generation, its efficiency increased steadily the later the generation in which it was imposed, and that this increase was greatest for characters of low heritability. They also took account, however, of the extra labour involved in carrying out this selection and came to the conclusion that single seed descent was the most cost-effective method of producing superior lines from a cross.

In the past, it has not been possible to obtain answers to the question of the efficiency of selection in the early generations of a pedigree by other than empirical methods, because the theory of response to selection in selfing programmes had remained undeveloped, unlike that concerning outbreeding populations of individuals (see, for example, Falconer, 1981). Very recently, however, Cornish (1988) has derived equations which describe the effects of selection during inbreeding on both the means and the genetical variance of subsequent generations. By simulating selection for characters with a range of heritabilities (0.50, 0.25, 0.10, 0.01` on the computer, it has, for the first time, been possible to find the optimum strategy for one-, two- and multi-stage selection. The results obtained so far show that, while optimum single stage selection in the F_2 or F_3 generation of the pedigree is less efficient than single-seed descent, that applied at F_4 or F_5 is very nearly as good, though never better than the latter. This is consistent with the empirical evidence from experiments on *N. rustica*. However, the response obtained from optimum multi-stage selection, where selection is practised in each of the four segregating generations of the pedigree, turns out to be approximately 20% better than single seed descent, which is, of course, a very worthwhile gain in efficiency. It must be emphasized, though, that to obtain this increase in efficiency, the breeder would need to raise the individuals of every segregating generation in a randomized experiment at agricultural density, and to score individual plants for the character in question. For most crops (but perhaps not rice), this requirement could be met only if the breeder was prepared to allocate more resources to the experiment than one involving single-seed descent.

This important theoretical investigation is still at an early stage and more need to be done before a clear picture emerges about the circumstances in which it is worthwhile applying selection in the early generations of a pedigree. If, however, this theory is supported by the results of actual breeding experiments, it should be possible to help breeders improve the efficiency of selection in their selfing programmes.

DISCUSSION

Earlier in this paper, we pointed out that, although there can be little doubt that the use of the pedigree method has been very successful, in that it has resulted in the introduction of hundreds of new and improved varieties, it cannot be regarded, from the genetical point of view, as very efficient. This is because, in practice, it appears that the majority of pedigrees are discarded and that only a very few turn out to be successful. So long as the breeder had no objective way of assessing the potential of a cross and, hence, of forecasting the value of the recombinant inbred lines that could be extracted from it, little could be done to improve the efficiency of the pedigree method. This situation has changed, however, because it is now possible to assess the value, of a cross by predicting the properties of these inbred lines from information obtained from the early generations of the pedigree. Provided that their distribution is approximately normal, it is a simple task to calculate the probability of obtaining a recombinant inbred line or a second cycle hybrid of the desired level of performance. If this probability turns out to be reasonably high, in terms of the number of lines that can be raised, the breeder will know, at the F_2 or F_3 stage of inbreeding, that the pedigree is worth retaining. If, on the other hand, this probability is unacceptably low, the cross may be discarded in the knowledge that nothing of value has been thrown away. In other words, the procedures described in this paper allow the breeder to recognize systematically and, hence, choose, the best crosses among the many made between the parental inbreds; this is the second of the three key stages of the pedigree method that we identified earlier.

The information required to predict the likely success of a cross demands, of course, a greater input of resources, for each cross fully assessed, than the conventional procedure. The breeder may feel, therefore, that the extra cost of these resources outweighs the value of the information obtained. There is no doubt that, in some crops, the cost of producing the seed of triple cross families is high, in terms of the requirement for skilled labour. It is worth pointing out, however, that in these circumstances these families could be replaced by a set of F_3 families, produced by selfing a random sample of F_2 individuals, from which an alternative estimate of *D* can be obtained (Jinks and Pooni, 1980). Though F_3 families are generally less informative than those of the triple test cross, if the breeder nevertheless decides to substitute the former for the latter, the only families that he would need to raise, over and above those raised with the conventional procedure, are the backcrosses, B_1 and B_2. Hence, the extra cost involved in raising the families required to provide the information needed to predict the parameters of the distribution of recombinant inbred lines that can be extracted from a cross need not be very great.

The individuals of these families must, however, be grown in the same randomized experiment, and care must be taken to ensure that the seed from which they are raised is of the same age and quality. The second requirement arises from the fact that, in many species, seed produced by emasculation and hand pollination is of a different quality from that produced by natural pollination. In barley, wheat and rice, for example, such seed is of lower quality than that produced naturally; hence, seedlings raised from it are less vigorous than those raised from self-set seed. In onions, on the other hand, the reverse is the case: seed from hand pollination is larger, and the seedlings and plants raised from it are better than those raised from natural seed. Because the seed of the F_1 and backcross generations has to be produced by hand, the seed of all families must be produced in this way. If this requirement is not met, the families of the basic generations and the triple test cross or F_3, could well turn out to be incapable of supplying the information that the breeder requires for prediction. The requirements that the families raised for this purpose should be grown in a single randomized experiment and that the seed from which they are produced should be of the same age and quality are perhaps, the most demanding. Accordingly they represent the largest additional costs that the breeder has to meet if he wishes to assess the value of his crosses in a systematic way.

Now it is obvious that the breeder is unlikely to be able to assess more than a small number of crosses fully. We may ask, therefore, what can be done by way of a preliminary assessment, so that attention can be concentrated on a handful of potentially the most promising crosses among the much larger number that have been made? The answer is that the best strategy is to choose those crosses of which the parents have a high average score for the character in question and of which the F_1 is also high scoring because it displays heterosis. The reason for the first requirement is simply that the breeder will be choosing parents with a high m. The reason for the second requirement is that, other things being equal, parental inbreds which differ for many genes that are dispersed between them and display dominance in the increasing direction are expected, when crossed, to give an F_1 which displays more heterosis than a hybrid from parents that differ for a smaller number of genes with the same properties. In other words, the magnitude of heterosis in the F_1 serves, in a relative sense, as a rough and ready measure of the number of genes for which the parents differ. In general, those crosses which segregate for a large number of genes are more valuable to the breeder, in terms of the transgressive variation they can engender, than those in which a smaller number of genes is segregating. Hence, if the breeder chooses high performing, heterotic crosses for full assessment, he will be choosing those in which his chance of success in extracting a superior recombinant inbred line or second cycle hybrid is greatest. In order to carry out a preliminary assessment of crosses, it would, of course, be necessary to raise each F_1 and its parental inbreds in a randomized experiment; however, because none of the families involved are segregating, this experiment, would not need to be very large.

There are two further points that are worth making about the procedures advocated in this paper. One of these concerns the third of the key stages of the pedigree method that requires critical evaluation, *i.e.* the efficient exploitation of the genetical variation for characters of interest to the breeder that is present in the F_2 family of a cross. The empirical evidence from experiments on *Nicotiana rustica* suggests that there are very few circumstances in which selection carried out during the course of inbreeding gives a better or quicker response than that obtained with single seed descent, unless the character in question has a very high heritability. The very recent theoretical investigation of the effect of selection, on the other hand, indicates that a worthwhile gain in response can be obtained, over a fairly wide range of heritabilities, if selection is practised in every generation. It must be emphasized, however, that this gain is obtained only with that combination of number and size of families which gives the optimum response to multi-stage selection; and that it is assumed that it is possible to score individuals raised in a randomized experiment in every generation during the course of inbreeding. In the light of this evidence, it is difficult to believe that the widespread practice whereby breeders attempt to select by eye for yield in the early generations of a pedigree, on material which is not randomized, can be justified on genetical grounds; it is thus unlikely to be very efficient. Hence, if the breeder is not able to apply the optimum strategy of multi-stage selection to his material, he would be better off deferring selection until the F_6 generation, as in single seed descent. On the other hand, the availability of this new theory and the opportunity it provides for finding the optimum strategy in any given situation, suggests that in the near future it may be possible to offer a service to the breeder which could advise him about the best strategy for his own particular circumstances.

The final point that we wish to make about the prediction procedures described in this paper concerns their validity and their application to situations of practical interest. As regards their validity, virtually all of the theory we have discussed has been thoroughly tested in experiments with *N. rustica*. The results obtained from these experiments show that, while predictions have nearly always turned out to be reliable, the actual proportion of lines or hybrids the performance of which equals or exceeds any given standard is usually greater than that predicted, even in crosses where it is known that complications like epistasis and linkage are present. These tests show, therefore that these prediction procedures are both robust and conservative. Of greater interest to the breeder, however, is that these procedures have been applied and continue to be applied to a wide range of crop species including barley (Thomas and Tapsell, 1983; Tapsell and Thomas, 1983), rice (Perera, Senadhira and Lawrence, 1985), onions (Werner *et al.*, 1988; Werner, Kearsey and Crowther, 1988) and the self-incompatible outbreeders, rye grass (Kearsey *et al.*, 1987; Devey *et al.* 1989) and brussels sprouts (Werner, Smith and Kearsey, 1986; Rogers, Kearsey and Smith, 1987; Werner *et al.*, 1989). The results from these investigations have shown that while, in general, it is relatively easy to obtain recombinant inbred lines that are superior for each character of interest to the breeder, it is much more difficult to extract lines that excel for each of these characters simultaneously — this problem will, of course, be very familiar to many experienced breeders. The range and volume of this work, however, leave little doubt that the procedure described in this paper are being actively applied to problems of practical importance in crops which span the full range of breeding systems and that there is no reason, of which we are aware, why they should not be applied to others, regardless of whether they are inbreeders or outbreeders

REFERENCES

BOUGHEY, H and JINKS, J L (1978). Joint selection for both extremes of mean performance and of sensitivity to a macro-environmental variable. III. The determinants of sensitivity, *Heredity, 40*, 363-369.

BRUMPTON, R J, BOUGHEY, H. and JINKS J L (1977). Joint selection for both extremes of mean performance and of sensitivity to a macro-environmental variable. I. Family selection, *Heredity*, *38*, 219-226.

CORNISH, M A (1988). Designing efficient, multistage selection, selfing programmes. *Proceedings of the Seventh Meeting of Eucarpia Biometrics Section*, As, Norway (in press).

DEVEY, F D, HAYWARD, M D, KEARSEY, M J, MCADAM, N J and EGGLESTON, M P (1989). Genetic analysis of production characters in *Lolium*. II. Triple test cross analysis of drill and plot performance. *Plant Breeding* (in press).

FALCONER D S (1981). *Introduction to Quantitative Genetics*. 2nd edition. Longman, London.

JINKS J L (1978). Unambiguous test for linkage of genes displaying non-allelic interactions for a metrical trait. *Heredity, 40,* 171-173.

JINKS J L (1981). The genetic framework of plant breeding. *Phil. Trans.R. Soc.* Lond. B 292, 407-419.

JINKS J L (1983). Biometrical genetics of heterosis. In Heterosis ed. R. Frankel. *Monograph on Theoretical and Applied Genetics*. Vol. 6, 1-45, Springer Verlag, Berlin.

JINKS, J L, JAYASEKARA, J E M and BOUGHEY, H (1977). Joint selection for both extremes of mean performance and sensitivity to a macro-environmental variable. II. Single seed descent. *Heredity, 39*, 345-355.

JINKS, J L and LAWRENCE, M J (1979). Current trends in biometrical genetics and their application to plant breeding. A review. *MARDI Res. Bull., 7*, 1-11.

JINKS, J L and LAWRENCE, M J (1983). The genetical basis of inbreeding depression and heterosis: its implications for plant and animal breeding. *PORIM Occasional Paper No. 11*, 1-8.

JINKS, J L and POONI, H S (1976). Predicting the properties of recombinant inbred lines derived by single seed descent. *Heredity, 36*, 253-266.

JINKS, J L and POONI, H S (1980). Comparing predictions of mean performance and environmental sensitivity of recombinant inbred lines based upon F_3 and triple test cross families. *Heredity, 45*, 305-312.

JINKS, J L and POONI, H S (1981a). Properties of pure breeding lines produced by dihaploidy, single seed descent and pedigree inbreeding. *Heredity, 46*, 391-395.

JINKS, J L and POONI, H S (1981b). Comparative results of selection in early and late stages of inbreeding. *Heredity, 46*, 1-7.

JINKS, J L and POONI, H S (1982). Predicting the properties of pure breeding lines extractable from a cross in the presence of linkage. *Heredity, 49*, 265-270.

KEARSEY, M J and JINKS, J L (1968). A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. *Heredity, 23*, 403-409.

KEARSEY, M J, HAYWARD, M D, DEVEY, F D, ARCIONI, S, EGGLESTON M P and EISSA, M M (1987). Genetic analysis of production characters in *Lolium*. I. Triple test cross analysis of spaced plant performance. *Theor. Appl. Genet., 75*, 66-75.

MATHER, K and JINKS, J L (1982). *Biometrical Genetics*. 3rd edition. Chapman and Hall, London.

PERERA, A L T, SENADHIRA, D and LAWRENCE, M J (1985). The genetical architecture of a number of economically important characters and the prediction of performance of recombinant inbred lines in rice. *International Rice Genetics Symposium*, 27-31 May 1985, IRRI, Manila.

PERKINS, J M and JINKS, J L (1970). Detection and estimation of genotype — environmental, linkage and epistatic components of variation for a metrical trait. *Heredity, 25*, 157-177.

POONI, H S and JINKS, J L (1976). The efficiency and optimal size of triple test cross designs for detecting epistatic variation. *Heredity, 36*, 215-227.

POONI, H S and JINKS, J L (1978). Predicting the properties of recombinant inbred lines derived by single seed descent for two or more characters simultaneously. *Heredity, 40*, 349-361.

POONI, H S and JINKS, J L (1979). Sources and biases of the predictors of the properties of recombinant inbreds produced by single seed descent. *Heredity, 42*, 41-48.

POONI, H S and JINKS, J L (1985a). Retrospective selection and sources of superior inbreds amongst pedigree inbred families of *Nicotiana rustica*. *Heredity, 54*, 255-260.

POONI, H S and JINKS, J L (1985b). Predicting the properties of first cycle inbreds and second cycle hybrids extractable from two, three and four parent crosses. *Heredity, 54*, 397-411.

POONI, H S, JINKS, J L and CORNISH, M A (1977). The causes and consequences of non-normality in predicting the properties of recombinant inbred lines. *Heredity, 38*, 329-338.

POONI, H S, JINKS, J L and de TOLEDO, J F F (1985). Predicting and observing the properties of second cycle hybrids using basic generations and inbred line x F_1 crosses. *Heredity, 54*, 121-129.

POONI, H S, JINKS, J L and YOHANNES, D (1985). Predicting the properties of random inbreds and second cycle hybrids using progeny families and hierarchical selfs of a four parent cross. *Heredity, 55*, 111-119.

POONI, H S, JINKS, J L and YOHANNES, D (1986). Comparison of inbred lines extracted from two single crosses and their double cross by single seed descent. *Heredity, 56*, 263-276.

ROGERS, W J, KEARSEY, M J and SMITH, B M (1987). The feasibility of producing inbred rather than F_1 hybrid cultivars in brussels sprouts: predictions from early generations. *Ann. Appl. Biol. III*, 677-688.

SIMMONDS, N (1979). Principles of Crop Improvement. Longman, London.

TAPSELL, C R and THOMAS, W T B (1983). Cross prediction studies on spring barley. II. Estimation of genetical and environmental control of yield and its component characters. *Theor. Appl. Genet., 64*, 353-358.

THOMAS, W T B and TAPSELL, C R (1983). Cross prediction studies on spring barley. I. Estimation of genetical and environmental control of morphological and maturity characters. *Theor. Appl. Genet., 64*, 345-352.

de TOLEDO, J F F, POONI, H S and JINKS, J L (1984a). Predicting the properties of second cycle hybrids produced by intercrossing random samples of recombinant inbred lines. *Heredity, 53*, 283-292.

de TOLEDO, J F F, POONI, H S and JINKS, J L (1984b). Predicting the transgressive potentials of *Nicotiana rustica* crosses for superior inbred lines and second cycle hybrids. *Vortr. Pflanzuchtg.*, 7, 151-166.

WERNER, C P, DOWKER, B D, DESOUZA, D C, SETTER, A P, CROWTHER, T C and HOROBIN J F (1988). Triple test cross predictions of the performance of recombinant inbred lines from a wide cross in onions (*Allium cepa* L.). *Ann. Appl. Biol., 112*, 525-535.

WERNER, C P, KEARSEY, M J, and CROWTHER, T C (1988). The prediction of inbred lines derived from a population cross in autumn-sown onions. *Proceedings of the Seventh Meeting of Eucarpia Biometrics Section*, As, Norway (in press).

WERNER, C P, SETTER, A P, SMITH, B M, KUBBA, J and KEARSEY, M J (1989). Performance of recombinant inbred lines in brussels sprouts (*Brassica oleracea* var. *gemmifera*). *Theor. Appl. Genet.* (in press).

WERNER, C P, SMITH, B M and KEARSEY, M J (1986). Tests of inbred predictions in brussels sprouts. *Proceedings of the Sixth Meeting of the Eucarpia Biometrics Section*, 125-139.