THE EXPECTED EFFICIENCIES OF HALF-SIB, TESTCROSS AND S1 PROGENY TESTING METHODS IN SINGLE POPULATION IMPROVEMENT

A. J. WRIGHT

Plant Breeding Institute, Cambridge

SUMMARY

Expressions are given for the expected rate of response to selection using half-sib, S1 or testcross progeny tests for biallelic loci with arbitrary dominance and epistatic properties. It is shown that the value of a tester is not necessarily expected to be a simple function of either its phenotype or of the number of recessive alleles it carries, and that neither downward selection nor inbreeding are expected to be consistently successful in isolating superior testers. Calculation of expected annual responses to selection under a variety of genetic models shows that, while mass selection is the most efficient method at high heritabilities, S1 testing is expected to be the best method for low heritabilities when replicated block trials are used. Conventional half-sib testing may have an advantage with complete randomisation of large families, especially for genetic models involving overdominance, deleterious recessives or epistasis in biennial or perennial crops. The use of the lowest homozygote tester was inferior to S1 testing for all models including simple directional dominance. Assumptions of low environmental homoeostasis of S1 material did not alter the general conclusions.

Monte Carlo simulations of breeding procedures confirmed the overall superiority of S1 testing in terms of response per cycle, and a system of mass selection followed by S1 testing gave a higher average response per year than did either mass selection or S1 testing alone.

1. INTRODUCTION

POPULATION improvement is the phase of a breeding programme in an outbreeding crop during which recurrent selection is applied to a variable population over a number of generations with the aim of improving both its performance and its potential to produce parents of superior synthetic or hybrid varieties. Although the simultaneous improvement of a pair of populations for the production of mutually adapted parents for hybrid varieties is sometimes used in maize (Comstock *et al.*, 1949; Russell and Eberhart, 1975), single populations are the rule for the majority of crops. Various methods of assessment and selection can be used, including individual phenotypic (mass) selection, full or half-sib family selection, and half-sib, test cross, or S1 progeny testing (where selfing is possible), and a choice will usually be made on grounds of both logistic convenience and expected efficiency.

In testcrossing, assessment and selection are based on progenies of the candidate genotypes with a common tester parent, whereas S1 progeny testing employs their first generation progeny derived by selfing. Although both these methods have been widely and successfully used in the improvement of maize populations, a theoretical basis for comparison of

their expected efficiencies with those of other methods of assessment and selection is lacking. This factor may be a deterrent to their more extensive application in other outbreeding crops which offer the necessary control of pollination.

For the majority of methods of assessment and selection used in population improvement, the expected genetic gains are functions of the population additive genetic variance. Such methods include mass phenotypic selection and all the common methods of family selection or progeny testing of full-sib families from balanced sets of crosses, or of half-sib groups of these families. The comparison of the expected efficiencies of these methods from a genetical point of view is therefore comparatively straightforward (Falconer, 1960), and depends only on the relative magnitudes of the various components of variance and not specifically on the mode of genetic control of the character.

This does not apply in general to either testcrossing or to S1 progeny testing. For both of these methods, the assessment values provide only indirect estimates of the true breeding values of the genotypes under test, so that the expected advance can only be quantified in terms of complex covariances which can involve cross products of additive, dominance and epistatic effects. Only with strictly additive gene action do these covariances reduce to simple functions of the additive variance.

Any theoretical examination of the expected efficiencies of these methods is therefore limited to the particular type of genetic control specified and convincing evidence as to the value of either method must be based on the examination of a range of such models. The same argument also applies to the use of Monte Carlo simulations of breeding procedures. This paper reports the results of an investigation which was designed to meet these requirements.

2. EXPECTED SELECTION RESPONSES

(i) *General formulation of response*

The most direct method of arriving at a formulation of the response following a single cycle of selection according to any system follows consideration of response as a correlated effect of selection for the criterion of assessment and utilisation of the known theory of correlated responses (Falconer, 1960). Thus:

$$R = \bar{\imath}\sigma_{A.Ax}/\sigma_x \tag{1}$$

where $\bar{\imath}$ is the intensity of selection in standard deviation units, $\sigma_{A.Ax}$ is the covariance of assessment values with the breeding or additive values of individuals and σ_x is the phenotypic standard deviation of assessment values containing genotypic and environmental components.

Further analysis requires the expansion of this expression in terms of a genetic model and this can be most easily done by first writing the statistics $\sigma_{A.Ax}$ and σ_x in terms of differential coefficients. It has been shown (Kojima, 1959; Wright, 1979) that any component of variance arising from a pair of biallelic loci A and B in linkage equilibrium and with allele frequencies p and u respectively $(q = 1 - p, v = 1 - u)$ can be expressed as the form

$$V_{rs} = p^r q^r u^s v^s (d^m y/dp^r \, du^s)^2 \Big/ \binom{n}{r}\binom{n}{s} (r! s!)^2 \tag{2}$$

where r is the order of interaction of alleles at locus A interacting epistatic-ally with the sth order at locus B and n is the level of ploidy. The term in brackets is the mth derivative $(m = r + s)$ of the population mean value (y) with respect to allele frequencies p and u. The variances associated with single locus effects are found by setting r or s to zero as appropriate. It can be similarly shown that the components of genetic covariance between any two sets of values corresponding to the population of genotypes can be written by replacing the squared derivative $(d^m y/dp^r\, du^s)^2$ in (2) by the product of the derivatives of the means of the two variates (say y and z), that is $(d^m y/dp^r\, du^s)\, (d^m z/dp^r\, du^s)$.

This method of formulation depends only on the binomial properties of the frequencies of genotypes in a population in Hardy–Weinberg equili-brium and the independence among loci specified by linkage equilibrium. It can therefore be applied to any set of assessment values associated with the population of individuals, provided that the mean of these values can be written in terms of the genotype frequencies at various loci. This allows a general formulation of the expected response following any method of assessment. Noting that permanent gains depend only on the additive component of covariance associated with the linear relationship between population mean and allele frequencies, then

$$R = \frac{\bar{\imath}\sigma_{A.Ax}}{\sigma_x} = \frac{\bar{\imath}\sum\limits_{i}^{k}\left[\frac{1}{2}pq(dx/dp)(dy/dp)+\frac{1}{2}uv(dx/du)(dy/du)\right]}{\sqrt{\sum\limits_{i}^{k}\sum\limits_{r}^{n}\sum\limits_{s}^{n}\left[p^r q^r u^s v^s (d^m x/dp^r\, du^s)^2 \Big/ \binom{n}{r}\binom{n}{s}(r!s!)^2\right] + V_{ex}}} \tag{3}$$

$$= \frac{\bar{\imath} W_{AyAx}}{\sqrt{V_{Ax} + V_{Dx} + V_{AAx} + V_{ADx} + V_{DAx} + V_{DDx} + V_{ex}}} \tag{4}$$

where x and y are the population means of assessment and genotypic values, and summation of r and s up to n gives all variance components for one pair of loci and summation is over k loci pairs. V_{ex} is a non-heritable component of variance among assessment values which depends on the type of assess-ment used.

Given an appropriate model to express the values of genotypes, the various derivatives of x and y can be expanded in terms of gene effects allowing a complete formulation of selection response for any system of assessment. The derivatives of y are a property only of the population under selection, while the testing system is specified in terms of x and its deriva-tives.

Genotypic values dependent on diploid biallelic epistatic loci can be expressed in terms of the model given by Mather and Jinks (1972) where d_a represents the effect of homozygosity at the first locus (A) and d_b that at the second (B), h_a and h_b represent the effects of heterozygosity at the two loci, i the interaction of homozygous states, l the interaction of heterozygous states, and j_a and j_b homozygous x heterozygous interactions where the suffix refers to the locus which is homozygous. The mean of the population is therefore

$$y = (p-q)d_a + (u-v)d_b + 2pq h_a + 2uv h_b + (p-q)(u-v)i$$
$$+ 2uv(p-q)j_a + 2pq(u-v)j_b + 4pquvl$$

and by differentiation

$$dy/dp = 2(d_a + (u-v)i + 2uvj_a + (q-p)(h_a + (u-v)j_b + 2uvl)) = 2\alpha_a,$$

where α_a is the population additive effect of locus A. Setting the parameters i, j_a, j_b, and 1 to zero for a non-epistatic locus, the more familiar form $dy/dp = 2(d_a + (q-p)h_a)$ is recovered. The eight derivatives of y with respect to p and u are given in table 1. A similar expansion of the mean of

TABLE 1

The eight derivatives of the population mean of genotypic values

dy/dp	$2(d_a + (u-v)i + 2uvj_a + (q-p)(h_a + (u-v)j_b + 2uvl))$
dy/du	$2(d_b + (p-q)i + 2pqj_b + (v-u)(h_b + (p-q)j_a + 2pql))$
d^2y/dp^2	$-4(h_a + (u-v)j_b + 2uvl)$
d^2y/du^2	$-4(h_b + (p-q)j_a + 2pql)$
$d^2y/dp\,du$	$4(i + (q-p)j_b + (v-u)j_a + (q-p)(v-u)l)$
$d^3y/dp^2\,du$	$-8(j_b + (v-u)l)$
$d^3y/dp\,du^2$	$-8(j_a + (q-p)l)$
$d^4y/dp^2\,du^2$	$16l$

assessment values in terms of this model allows the evaluation of derivatives and expression of the expectation of response for any method of assessment and selection.

(ii) *Testcross systems*

With all forms of testcrossing, the progenies used for assessment are generated by crossing the parental population with a fixed tester which may be a single genotype or a population, commonly the parental population itself (conventional half-sib or general combining ability selection). In general, for a tester with allele frequencies p' and u' at A and B, the mean value of all progenies can be written as

$$x = sd_a + td_b + wh_a + zh_b + sti + szj_a + twj_b + wzl,$$

where $s = (pp' - qq')$, $t = (uu' - vv')$, $w = (p'q + pq')$, and $z = (u'v + uv')$. Then

$$dx/dp = d_a + (uu' - vv')i + (u'v + uv')j_a$$
$$+ (q' - p')(h_a + (uu' - vv')j_b + (u'v + uv')l) = \alpha_{at} \tag{5}$$

which is a generalisation of the term given by Rawlings and Thompson (1962). Differentiation with respect to u gives a similar term for locus B, α_{bt}. Further differentiation leads to

$$d^2x/dp.du = i + (v' - u')j_a + (q' - p')j_b + (v' - u')(q' - p')l = (\alpha\alpha)_t,$$

and

$$d^2x/dp^2 = d^2x/du^2 = d^3x/dp^2\,du = d^3x/dp\,du^2 = d^4x/dp^2\,du^2 = 0,$$

because the progeny means show variance due to additive and additive $\times$ additive effects only. When the parental population is used as tester, then $p' = p$ and $u' = u$, so that $\alpha_{at} = \alpha_a$ and $(\alpha\alpha)_t = (\alpha\alpha)$. The expectations of α_{at}

and $(\alpha\alpha)_t$ for various tester genotypes and types of gene action are given in table 2.

Expectations of α_{at} (additive) and $(\alpha\alpha)_t$ (additive × additive) gene effects for different tester genotypes and types of gene action

Effect	Tester genotype	General form	Complementary epistasis	Duplicate epistasis
α_{at}	aabb	$d_a - vi + uj_a + h_a - vj_b + ul$	$4u$	$4v$
	aaBB	$d_a + ui + vj_a + h_a + uj_b + vl$	4	0
	AAbb	$d_a - vi + uj_a - h_a + vj_b - ul$	0	0
	AABB	$d_a + ui + vj_a - h_a - uj_b - vl$	0	0
$(\alpha\alpha)_t$	aabb	$i + j_a + j_b + l$	4	-4
	aaBB	$i - j_a + j_b - l$	0	0
	AAbb	$i + j_a - j_b - l$	0	0
	AABB	$i - j_a - j_b + l$	0	0

The variance of assessment values σ_t^2 contains an additional component V_{et} which is a complex involving environmental variation among progenies, and genetic and environmental variation arising within progenies. With large progenies the last two sources can be ignored. For the purposes of the arguments which follow, it is assumed that V_{et} is constant for all testers. The expected response to one cycle of testcross evaluation and selection can therefore be written as

$$R_t = \bar{\imath}\sigma_{A.At}/\sigma_t = \bar{\imath}\frac{\sum(pq\alpha_a\alpha_{at} + uv\alpha_b\alpha_{bt})}{\sqrt{\sum(\frac{1}{2}pq\alpha_{at}^2 + \frac{1}{2}uv\alpha_{bt}^2 + \frac{1}{4}pquv(\alpha\alpha)_t^2) + V_{et}}} \tag{6}$$

where summation is over all pairs of interacting loci.

(a) *The genotype of the optimum tester*

For a population with a single locus segregating, it is clear that the greatest response will follow assessment with a tester for which α_{at} is maximised, as $\sigma_{A.At}$ and the genetic portion of σ_x are increased proportionately, thus reducing the influence of V_{et}. Since the values of α_{at} for such a locus are $(d+h)$, d, and $(d-h)$ for the tester genotypes aa, Aa, and AA respectively (where A is the dominant allele), the homozygous recessive is clearly the best tester unless α_a is negative due to overdominance with p in excess of its optimum. This last exception is well known, but others may arise when a pair of epistatic loci is considered (table 2). With complementary interaction, tester $aaBB$ gives a higher α_{at} value than does $aabb$, but a zero α_{bt}. For many combinations of allele frequencies p and u, one or other of the dispersed homozygotes is likely to be the most effective tester parent.

In systems with many loci, the superiority of the completely recessive tester cannot be assured even in the absence of epistasis, although this has frequently been assumed. Although an increase in α_{at} will raise the response from any specific locus (A), the extra variance among testcross means serves only as an additional source of error for all other loci, so that the total response will not necessarily increase. The covariance $\sum pq\alpha\alpha_t$ can also be regarded as the sum of products across loci of α and α_t weighted by

the coefficients pq. Thus the maximum response is to be gained not simply from a tester generating the maximum overall α_t values, but those which are also completely correlated with α, as this has the property of distributing the selection differential as efficiently as possible with respect to the potential of each locus to respond. The use of the completely homozygous recessive tester on the one hand and conventional half-sib testing on the other hand represent alternative strategies, the first maximising the total selection differential, the second ensuring its optimal utilisation. No generalisation can therefore be made as to the superiority of either of these or any other tester genotypes, except that the smaller is V_{et} (i.e., the higher the broad-sense heritability), the higher is the relative efficiency of conventional half-sibs likely to be.

(b) *The choice of tester by phenotypic selection*

Even in the circumstances where the completely recessive homozygote is expected to be the best tester, there is no way of obtaining such a genotype with respect to any character with complex inheritance. On the assumption that dominance is directional and positive for most loci, the isolation of superior testers by reduction of gene frequencies through selection for low performance has been suggested by Allison and Curnow (1966) and Moreno-Gonzalez and Grossman (1976). Since the purpose of applying testcrossing procedures is normally to deal with a character for which phenotypic mass selection is ineffective, such a procedure would not always be expected to succeed.

The change in tester efficiency can be treated exactly as any other character under selection. If the starting point for isolation of a tester or testers is the parental population itself, then one cycle of low phenotypic selection in the case of no epistasis leads to a change of gene frequency of

$$\Delta_{p'} = -\bar{\imath}2pq\alpha_a/\sigma p \qquad (7)$$

where σ_p is the phenotypic standard deviation. The rate of change of tester efficiency with changing gene frequency is found by differentiation of (6) with respect to p' as

$$dR/dp' = -\bar{\imath}(4pqh_a\alpha_a)V_{et}/\sigma_x{}^3 \qquad (8)$$

so that the product of (7) and (8) summed over loci gives the change in efficiency. Expressed as a proportion of the response to be gained from the use of the unselected parental population itself as tester (conventional half-sibs) this is

$$\frac{\Delta R}{R} = \frac{\bar{\imath}V_{et}}{\sigma_p\sigma_x^2}4(\sum p^2q^2h_a\alpha_a^2)/(\sum pq\alpha_a^2) \qquad (9)$$

where $\bar{\imath}$ and σ_p relate to the selection of the tester and V_{et} and σ_x^2 to the testcrossing phase itself. The ratio depends to a large extent on the proportion of testcross variance which is environmental (i.e., V_{et}/σ_x^2) because V_{et} increases the importance of efficient tester genotype. When V_{et} is zero, any selection away from the parental population itself would be detrimental. Otherwise it is clear that the likely gain from downward selection for testers is positive when dominance for all loci is strong and positive, particularly for those with large amounts of additive variance. As

expected, it will be ineffective when dominance is ambi-directional. The necessary correlation between dominance effect and additive variance will be promoted by the presence of loci with grossly unequal effects overall, but would otherwise be expected to occur only by chance in a breeding population resulting from crossing unrelated material. This suggests that short-term phenotypic selection is unlikely to be consistently successful in effecting large improvements in tester efficiency, and similar selection on the basis of a progeny test hardly better.

It has to be remembered that ineffective selection will in fact be detrimental, as the expected efficiency of a tester genotype chosen at random is lower than that of the population as a whole. Moreno-Gonzalez and Grossman (1976) concluded that low selection would be effective in isolating superior testers for use in reciprocal selection schemes, but since their arguments were based on single loci, this outcome remains in doubt. Although formula (8) is not strictly valid for the choice as testers of varieties or lines whose performance is already well known, the same argument as to the necessity for a correlation of additive and dominance effects holds.

As an alternative to selection, inbreeding has been suggested as a means of isolating efficient testers (Cress, 1967; Russell and Eberhart, 1975). However, the efficiency of a tester depends linearly on its gene frequencies irrespective of their degree of homozygosity. Expansion of (5) for genotypes with different levels of heterozygosity shows that the average value of α_{at} is unaffected by epistasis of unlinked genes provided the average allele frequency is held constant. Inbreeding will increase the variance of tester efficiency as that of any other character, however, thus enlarging the scope for selection. The expected change in tester efficiency for one cycle of low selection among fully inbred individuals in the absence of epistasis, again as a proportion of the efficiency of conventional half-sibs, is

$$\frac{\Delta R}{R} = \frac{\bar{i} V_{et}}{\sigma_p \sigma_x^2} \cdot 8(\sum p^2 q^2 h_a d_a \alpha_a)/(\sum pq\alpha_a^2).$$

The additional factor of two when compared with (9) reflects the effects of increased variation, but otherwise the success of this selection of testers depends on a correlation of α_a, d_a, and h_a terms.

(iii) *S1 testing*

The mean genotypic value of any progeny produced by one generation of selfing can be derived from that of its parent by applying coefficients of $\frac{1}{2}$ to all h and j terms and $\frac{1}{4}$ to all l terms. The expected response to selection based on S1 progeny means is therefore a simple modification of that for mass phenotypic selection:

$$R_s = \bar{i}\sigma_{A_yA_s}/\sigma_s = \bar{i}W_{A_yA_s}/\sqrt{V_{As} + V_{Ds} + V_{AAs} + V_{ADs} + V_{DAs} + V_{DDs} + V_{es}}$$

where

$$W_{A_yA_s} = \sum (2pq\alpha_a\alpha_{as} + 2uv\alpha_b\alpha_{bs})$$

and

$$V_{A_s} = \sum (2pq\alpha_{as}^2 + 2uv\alpha_{bs}^2), \text{ etc.}$$

and V_{es} is the error variance of an S1 progeny mean due to non-heritable and within family sources. All effects with an s suffix are obtained by making the above substitutions to the usual genotypic effects as given in table 1. In particular,

$$dx/dp = 2\alpha_{as} = 2(d_a + (u - v)i + uvj_a + \tfrac{1}{2}(q - p)(h_a + (u - v)j_b + uvl)).$$

As for testcrossing, response depends on a weighted sum of products across loci, this time of α and α_s, but in this case its magnitude depends only on the genetic properties of the population and not on any choices open to the breeder. The parameters α and α_s are equal only with additive gene action, but their correlation will be high when loci show similar homozygote, heterozygote and epistatic properties (d, h, i, j and l). For example, with complete positive dominance at loci with equal effects, the correlation will be perfect. Although non-additive effects therefore need not necessarily lead to a reduced correlation, in practice they will make it more probable, particularly when overdominance or strong epistasis is operative. These are also the conditions when inbreeding depression will occur.

Major deleterious recessive genes are of particular interest, as their homozygotes are known to contribute to the depression of S1 progenies. For a rare recessive allele $\alpha_a \approx 0$, whereas $\alpha_{as} \approx \tfrac{1}{2}d_a$, so that some selective effort will be wasted on these loci. However, the influence of such loci will be small unless they are numerous or of very large effect, as the variance generated is also a function of the product pq.

A further consideration in the case of S1 progenies is the possibility of lowered environmental homeostasis leading to higher error variances and reduced selection responses. Such lack of homoeostasis can be expected for characters which show inbreeding depression in terms of mean expression, such as yield itself. The magnitude of this effect and its influence on selection efficiency is difficult to predict, but for first generation inbreds may not be serious in many crops.

3. COMPARISONS OF EXPECTED GAINS FOR SOME GENETIC MODELS

Because the statistics determining response to selection are specific to the method employed, comparisons of expected responses can only be made after evaluation of these statistics for particular genetic situations. The expected gains per cycle were calculated in this way for mass selection, half-sib progeny testing, testcross progeny of the lowest completely homozygous tester and S1 testing, for eight different genetic models. In each model, 12 loci were either all given equal single gene and epistatic effects, or were divided into two equal groups with different effects, generating seven combinations of mono- and ambi-directional dominance and overdominance and duplicate and complementary epistasis. Such over-dominance may represent pseudo-overdominance of chromosome segments which maintain their integrity through single cycles of selection. Ten representative sub-models of each type were produced by the random generation of the allele frequency at each locus in the range 0 to 1. In the eighth model, six loci were designated as major deleterious recessives with allele frequencies fixed at $0{\cdot}95$, each with four times as great an effect for all ten sub-models than the other loci which had a mixture of dominance properties. The values of $W_{A.A_s}$, V_A, and σ^2_{Gx} and σ^2_G (total genetic variance

among progeny means and population genotypes respectively) were computed for each method using the formulae given earlier, and expected gains found assuming various narrow-sense heritability levels. Since the heritable properties of the models were fixed, heritability was adjusted by varying the non-heritable variance (V_e), and for some models which generated large amounts of non-additive variance, the higher levels were unobtainable. Complete randomisation of n individual plants was initially assumed, and the within family sources of variation ignored so that V_{ex} was equal to V_e/n for all progeny testing methods.

Gains per year are also dependent on the cycle length of the method used. For most crops, mass selection requires one season for intercrossing in addition to the period of assessment which depends on the life cycle and the agronomy of the species; progeny tests require an additional season for the production of progeny. Gains from progeny testing and mass selection can be put on the same footing by multiplication of the latter by c, the ratio of the number of years per cycle for the two systems. This takes values of $1\cdot5$ and $1\cdot33$ for annual and biennial crops respectively. In principle, c could be increased further to take account of the higher selection intensities likely with mass selection.

As expected, increases in the error variance led to reductions in expected gain from all methods at all family sizes, mass selection being the most seriously affected and S1 testing the least. For all models, S1 testing was predicted to be the most effective method at low heritability and mass selection at high heritability, with half-sibs showing some superiority at intermediate values in some circumstances. Testcrossing was competitive with the other methods only with the directional dominance model. These patterns of changing relative responses can be effectively described by determining the portions of the heritability spectrum in which one particular method is expected to be better than another. The critical heritability value below which a particular progeny test gives higher gain than mass selection can be computed as

$$h^2 = \frac{n(W_{A_yA_x})^2 - c^2(V_A)^2}{c^2 V_A(n\sigma_{Gx}^2 - \sigma_G^2)}.$$

When gene action is strictly additive, then $W_{A_yA_x}$ is a multiple of V_A for all methods, and the formula simplifies in the case of half-sib progeny testing to

$$h^2_{HS} = (n - 4c^2)/c^2(n - 4),$$

which is at a minimum because its usual advantage of freedom from non-additive effects is lost. For S1 testing, on the other hand, the critical heritability with an additive model is a maximum, when

$$h^2_{(S1)} = (n - c^2)/c^2(n - 1).$$

When n is very large, both these formulae tend towards a limit of $1/c^2$, which is $0\cdot44$ and $0\cdot59$ for annuals and biennials respectively.

Critical heritabilities below which the different progeny testing methods are expected to be better than mass selection for different values for n and c (and assuming equal selection intensity) are expressed in table 3a as the mean of the values for the ten submodels. Those below which S1 testing exceeds half-sib testing appear in table 3b, but these are independent of c.

TABLE 3

Critical heritability values (per cent)

				(a) Values below which the progeny test gives higher gains per year than mass selection					(b) Values below which S1 exceeds half-sib testing		
		HS		S1		TX					
Model	c	n20	5	20	5	20	5	20	5	2	
Additive	1·5	31	0	42	31	31	0	100	100	100	
	1·3	45	0	54	45	45	0	(100)		(100)	
Directional	1·5	34	0	36	26	30	20	41	62	69	
dominance	1·3	50	0	47	40	41	34	(32)		(66)	
Ambidirectional	1·5	33	0	36	25	8	0	42	65	72	
dominance	1·3	50	0	47	39	13	0	(31)		68)	
Directional	1·5	43	0	27	15	24	16	16	28	34	
overdominance	1·3	46	0	36	29	31	24	(9)		(30)	
Ambidirectional	1·5	42	0	27	15	0	0	16	31	39	
overdominance	1·3	46	0	36	28	0	0	(8)		(31)	
Deleterious	1·5	36	0	24	19	6	2	16	36	49	
recessives	1·3	53	0	32	27	7	4	(13)		(45)	
Duplicate	1·5	35	0	31	25	30	25	23	37	42	
epistasis	1·3	49	0	38	36	38	34	(19)		(41)	
Complementary	1·5	34	0	31	15	26	9	26	44	52	
epistasis	1·3	51	0	41	31	36	20	(13)		(44)	

The standard errors of all these means are very small (typically of the order of 0·02). It is clear from comparisons of these values that for almost all heritability values for all models, either mass selection or S1 testing is expected to be the best method. The only exceptions are for the higher family size ($n = 20$) where half-sib testing has a small range of superiority to S1 testing for both annual and biennial crops for the models involving overdominance, epistasis, or major recessive genes. The critical values for testcrossing have not been so extensively explored as for the other models, as these are lower than for S1's for all models and family sizes, and only exceed those for half-sibs for small families.

The advantage of half-sibs with large families can be further examined by assuming extremely large n values. Half-sib testing now exceeds mass selection for all heritabilities below $1/c^2$, that is 0·44 for annuals and 0·57 for biennials. Since the critical heritability values for S1 testing are barely increased above those for $n = 20$, half-sib testing is always expected to be the best method for large completely randomised families. In practice, however, this is an unlikely situation as large families will usually be grown in replicated blocks, so that the error variance of a family mean V_{ex}, is the compound $(V_b/r + V_e/n)$ where V_b is the error variance among plots. For large families, this is now highly dependent on r, the number of replicate plots. Although comparisons with mass selection are more difficult in these circumstances, the relevant comparisons among progeny testing methods can be made by replacing n in formulae for critical heritability by r. Thus for typical numbers of replicates (2 to 5) table 3b shows that S1 is strongly favoured for models with simple additive or dominance properties, and for all models over the range of heritabilities in which mass selection would not be preferred.

The possibility of reduced homoeostasis of S1 progenies can be allowed for by assuming a family size smaller than that actually grown when calculating responses or critical heritabilities. Thus the values in brackets for n (or r) values of 20 and 2 in table 3b assume that the error variance per S1 plant is double that for a member of a half-sib family. Although this would add significantly to the superiority of half-sibs with large families, for low values of n or r the critical heritabilities are little changed.

In an attempt to attain an even more comprehensive treatment of the possible range of models of genetic control, a further study was carried out in which models were derived by the generation of additive and dominance effects at random within a specified range so as to give a complex mixture of different types of control, allowing the inclusion in turn of mono- and ambi-directional dominance and overdominance. Although the sub-models showed a greater range of variability, the mean critical heritabilities were sufficiently similar to those of their counterparts with equal gene effects to leave the existing conclusions unaffected.

A clear pattern emerges from the results of this section. While mass selection gives the greatest gains at high heritability values, S1 testing is the best method for lower values in typical replicated blocks trials or for completely randomised small families. With completely randomised large families, the situation giving the greatest gain of all, half-sibs may be favoured, particularly at intermediate heritabilities arising from complex genetic models in biennial or perennial species. Testcrossing is only competitive for simple dominance models and small family or replicate numbers, but never exceeds S1 testing.

4. SIMULATIONS OF BREEDING PROGRAMMES

Evaluation of the formulae for expected response to selection according to a variety of genetic models has led to the prediction that S1 testing is likely to be generally superior to testcrossing or half-sib progeny testing. In an attempt to strengthen the basis for this prediction, Monte Carlo type simulations of breeding procedures were carried out. All simulations assumed the selection of 10 individuals from a population of 100 at each cycle, with progenies of 20. The models used assumed 18 unlinked loci initially in gametic equilibrium with beneficial alleles at a frequency of 0·3. The forms of gene action chosen were those expected to be the most unfavourable to S1 testing and involved directional and ambi-directional complete and overdominance in combination with various forms of epistasis. After calculation of the components of genetic variance according to the model, the error variance was adjusted to give a chosen value for the initial heritability assuming a completely randomised design, and phenotypes produced using error deviates generated at random according to this distribution throughout. Two studies were carried out.

(i) *Comparison of mass selection, half-sib, testcross and S1 progeny testing*

Four models (A, B, C and D) giving combinations of ambi-directional dominance and overdominance with and without duplicate epistasis, and a fifth (E) combining duplicate epistasis at 10 loci with complementary epistasis at 8, were each run four times, twice for each initial heritability

value of 0·4 and 0·2. The tester parent was assumed to be the lowest homozygote, completely recessive in the case of complementary epitasis. The response to selection after four cycles was subjected to analysis of variance (table 4). The large mean square for models was largely due to the

TABLE 4

Mean responses after four cycles of selection according to four methods and under five genetic models and their analysis of variance

	Mass	S1	Hs	Tx
A	15·0	22·0	17·2	12·9
B	7·1	11·3	7·6	5·4
C	15·1	22·9	16·6	12·9
D	7·9	10·3	8·7	5·6
E	20·8	21·8	19·9	21·5
Mean	13·2	17·7	14·0	11·7

Source	df	Mean square
1. Methods	3	129·8***
(a) S1 *vs* others	1	333·2***
(b) Remainder	2	28·0*
2. Models	4	547·9***
3. Heritabilities	1	23·2*
4. Me × Mo	12	12·0
(a) S1 *vs* Tx *vs* others × E *vs* others	2	40·6**
(b) Remainder	10	6.2
5. Me × H	3	4·2
6. Mo × H	4	7·4
7. Me × Mo × H	12	1·8
8. Duplication	40	6·2

*** signifies $P < 0·001$, ** $P < 0·01$, * $P < 0·05$

different way the genetic effects had to be specified, and so is of little meaning. The significant methods item was almost entirely attributable to the superiority of S1 testing over all models and heritability levels, with an average response 50 per cent greater than that for the other two methods for models A, B, C and D and 5 per cent greater than for E. Testcrossing was the worst method for all models except E, where it equalled half-sib and mass selection. As expected, mass selection was the method most affected by the initial heritability, but this difference was not significant. In terms of response per annum, mass selection would certainly have exceeded half-sib testing throughout and approached S1 testing for the mixed epistatic model (E), especially if allowance was made for the higher selection intensity likely.

(ii) *Comparisons of breeding programmes*

In this study, selection was placed in the context of a critical breeding programme by assuming that initial mass selection would be replaced by S1 testing as soon as heritability fell from an initial value of 0·5 to the critical value estimated for that particular model for an annual crop (table 3). This system was compared with simple mass selection and S1 testing over $(m+6)$ years, where the switch occurred at m years. Thus the mass selection

programme had completed $(m+6)/2$ cycles at the time of assessment, the S1 programme $(m+6)/3$ and the combined system $(m+4)/2$. The five models used (A, B, C, D and E) were complete directional and ambi-directional dominance, overdominance, and duplicate and complementary epistasis respectively.

Since the evaluation of the three methods takes place after a certain number of years which is specific to each particular replicate run, the analysis of variance takes a form analogous to a split-plot design (table 5). This

TABLE 5

Mean responses to selection according to three methods evaluated after a fixed number of years and their analysis of variance (see text)

	Mass	S1	Combined
A	27·6	26·0	29·5
B	24·2	20·7	22·9
C	15·8	17·9	17·4
D	52·2	50·5	53·8
E	34·8	36·8	36·5
Mean	30·9	30·4	32·0

Source	df	Mean square
1. Models	4	2242·0***
2. Reps within models	15	15·4
3. Methods	2	14·1*
(a) combined *vs* others	1	25·4**
(b) Remainder	1	2·8
4. Me × Mo	8	8·0*
(a) Mass *vs* others × B *vs* others	1	24·2**
(b) Remainder	7	5·7
5. Me × Reps (Error)	30	3·2

*** signifies $P<0.001$, ** $P<0.01$, * $P<0.05$

shows that the combined method gave significantly more advance (5 per cent) than the other two methods overall, but that there was a significant methods × models interaction. This was largely due to the superiority of mass selection for model B (ambi-directional dominance), and was unexpected since known differences among the models were expected to have been accounted for by the use of appropriate critical heritability values.

5. DISCUSSION

An examination of the expected rate of response to selection according to any method requires consideration of both the numerator $(\sigma_{A.Ax})$ and the denominator (σ_x) of the response formula. In the case of testcross systems, examination of σ_x alone is likely to lead to the conclusion that the completely recessive homozygote is the best tester (Hull, 1947) as it maximises the total genetic variance among progenies, whereas consideration only of $\sigma_{A.Ax}$ would suggest the use of the parental population itself. This latter conclusion was arrived at by Allison and Curnow (1966) when

comparing varietal testers for which σ_x could be assumed constant. Expansion of both these statistics in terms of many genes has shown here that the efficiency of a tester for the improvement of a quantitative trait is not in general expected to be a simple function either of its phenotype or of the number of recessive alleles it carries. When the choice of tester is based on its phenotype, whether this choice is confined to genotypes drawn from the parental population (Allison and Curnow, 1966) or are unrelated varieties of lines (Rawlings and Thompson, 1962) the effectiveness of this choice depends on a correlation between additive (α) and dominance (h) effects across loci, and cannot generally be expected to be successful. Some forms of epistasis have also been shown to have a detrimental effect on the correlation between phentoype and tester efficiency. In spite of this prediction, however, there is some evidence that tester efficiency is associated with low yield among lines of maize (Rawlings and Thompson, 1962; Lonnquist and Lindsay, 1970).

Inbreeding *per se* is not expected to have any effect on mean tester efficiency as it does not cause changes in gene frequency. The observed advantage of inbred testers in maize (Horner *et al.*, 1963, 1973) is only explicable in terms of different gene frequencies from those in the outbred testers with which they were compared.

Apart from some formulae based on segregation at a single locus (Wricke, 1976) information on the efficiency of S1 testing comes from experiments on maize and from Monte Carlo simulations. Choo and Kannenberg (1979a) used the latter method to demonstrate higher gains per cycle for S1 testing than either mass or modified ear-to-row selection as used in maize (intercrossing of selected half-sib families), but only for simple additive and directional dominance models. Wright (1977) found that S1 testing compared favourably with half-sib and testcrossing for a range of models which included ambi-directional dominance and epistasis, and the present study has confirmed that this method has a robustness in the face of differing genetic conditions which lead to its general superiority over other progeny testing methods for typical trial designs. The considerable weight of experimental evidence in maize also indicates a general superiority of this method (Genter and Alexander, 1966; Genter, 1973; Carangal *et al* 1971; Lonnquist, 1968).

Some workers have drawn attention to the higher error and genotype × environmental interaction variances inherent in the use of inbred material. Their data show approximately twofold increases for both of these components in the case of rye (Wricke, 1976) and threefold for genotype × year interactions in maize (Lonnquist and Lindsey, 1964). The present work has shown that increases of this order are unlikely to influence the choice of progeny test.

The limited size of selected sample used in most practical breeding programmes means that alleles which are at low frequency and under low selection pressure are in danger of loss. In their simulation studies, Choo and Kannenberg (1979b) found that S1 testing led to the loss of more desirable alleles than did mass selection. No such differential was recognised in the present experiment, either in terms of alleles lost at any stage or of the population mean after many generations (20) of selection, but this may not be surprising in view of the higher initial allele frequencies assumed (0·3 c.f. 0·1). The selection pressures applied to any allele by half-sib testing

and the use of a low homozygous tester are $(\bar{i}pq\,\alpha_a/\sigma_{hs})$ and $(\bar{i}pq\,\alpha_{at}/\sigma_t)$ respectively where σ_{hs} and σ_t are the phenotypic standard deviations of progeny means in each case. For a rare allele, $\alpha_a \approx (d+h) = \alpha_{at}$, so that the ratio of selection pressures for the two methods is close to σ_t/σ_{hs}, which may be much larger than unity. This advantage of half-sib testing applies similarly to mass selection, but S1 testing would be expected to be inter-mediate in its properties. Using the same argument, the completely reces-sive tester would have the same disadvantage for rare dominants but an advantage with respect to the conservation of rare recessives.

For annual crops such as maize, S1 progenies serve the dual purpose of assessment criterion and means of maintenance of genotypes, in contrast to testcross progenies which generally have no inherent value and half-sibs which offer only one-half of the gain when used in preference to their parents. Breeders of crops which can be clonally propagated have a choice of material for selection, the gains offered being unaffected unless different degrees of linkage disequilibrium are generated.

In conclusion, evidence from both theoretical and simulation studies support the use of S1 progeny testing for most genetic and experimental situations where heritability is too low for efficient use of mass phenotypic selection. Its successful use also depends on the ease with which selfed seed can be obtained. Unlike maize, many cross fertilising crop species have self incompatibility systems, but in many cases these allow some seed to be set under selfing, and in others the incompatibility mechanism can be overcome by means of bud pollination, carbon dioxide atmospheric enrichment, or high temperatures, and this is commonly practised in the production of inbred lines. The theoretrical justification of S1 testing together with its success in maize may encourage breeders of these crops to adopt S1's as a means of progeny testing.

6. REFERENCES

ALLISON, J. C. S., AND CURNOW, R. N. 1966. On the choice of tester parent for the breeding of synthetic varieties of maize. *Crop Sci., 6*, 541-544.

CARANGAL, V. R., ALI, S. M., KOBLE, A. F., RINKE, E. H., AND SENTZ, J. C. 1971. Comparison of S1 with testcross evaluation for recurrent selection in maize. *Crop Sci., 11*, 658-661.

CHOO, T. M., AND KANNENBERG, L. W. 1979a. Relative efficiencies of population improve-ment methods in corn: a simulation study. *Crop Sci., 19*, 179-185.

CHOO, T. M., AND KANNENBERG, L. W. 1979b. Changes in gene frequency during mass, modified ear-to-row, and S1 selection: a simulation study. *Crop Sci., 19*, 503-509.

COMSTOCK, R. E., ROBINSON, H. F., AND HARVEY, P. H. 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. *Agronomy Journal, 41*, 360-367.

CRESS, C. E. 1967. Reciprocal recurrent selection and modifications in simulated populations. *Crop Sci., 7*, 561-567.

FALCONER, D. S. 1960. *Introduction to Quantitative genetics.* Oliver & Boyd, London.

GENTER, C. F. 1973. Comparison of S1 and testcross evaluation after two cycles of recurrent selection in maize. *Crop Sci., 13*, 524-527.

GENTER, C. F., AND ALEXANDER, M. W. 1966. Development and selection of productive S1 inbred lines of corn. *Crop Sci., 6*, 429-431.

HORNER, E. S., LUNDY, H. W., LUTRICK, M. C., AND WALLACE, R. W. 1963. Relative efficiencies of recurrent selection for specific and for general combining ability in corn. *Crop Sci., 3*, 63-66.

HORNER, E. S., LUNDY, H. W., LUTRICK, M. C., AND CHAPMAN, W. H. 1973. Comparison of three methods of recurrent selection in maize. *Crop Sci., 13*, 485-489.

HULL, F. H. 1947. Cryptic homozygous lines. *J. Am. Soc. Agron., 39,* 438-439.

KOJIMA, K. 1959. Role of epistasis and overdominance in stability of equilibria with selection. *U.S. National Academy of Sciences, 45,* 984-989.

LONNQUIST, J. H. 1968. Further evidence on testcross versus line performance in maize. *Crop Sci., 8,* 50-53.

LONNQUIST, J. H., AND LINDSAY, M. F. 1964. Topcross versus S1 line performance in corn. *Crop Sci., 4,* 580-584.

LONNQUIST, H. H., AND LINDSAY, M. F. 1970. Tester performance level for the evaluation of lines for hybrid performance. *Crop Sci., 10,* 602-604.

MATHER, K., AND JINKS, J. L. 1972. *Biometrical Genetics, 2nd ed.* Chapman & Hall, London.

MORENO-GONZALEZ, J., AND GROSSMAN, M. 1976. Theoretical modification of reciprocal recurrent selection. *Genetics, 84,* 95-111.

RAWLINGS, J. O., AND THOMPSON, D. L. 1962. Performance level as a criterion for the choice of maize testers. *Crop Sci., 2,* 217-220.

RUSSELL, W. A., AND EBERHART, S. A. 1975. Hybrid performance of selected maize lines from reciprocal recurrent and testcross selection programmes. *Crop Sci., 15,* 1-4.

WRICKE, G. 1976. Comparison of selection based on yield of half-sib progenies and of 11 lines *per se* in rye. *Theoretical and Applied Genetics, 47,* 265-269.

WRIGHT, A. J. 1977. Annual Report of Plant Breeding Institute, Cambridge, England.

WRIGHT, A. J. 1979. The use of differential coefficients in the development and interpretation of quantitative genetic models. *Heredity, 43,* 1-8.