Theor Appl Genet (1993) 87:152–160

THEORETICAL AND APPLIED GENETICS

The effect of gametic-phase disequilibrium on the prediction of response to recurrent selection in plants

I. J. Mackay[1], **J. P. Gibson**[2]

[1] Lion Seeds Ltd., Woodham Mortimer, Maldon, Essex CM9 6SN, England, UK
[2] Centre for Genetic Improvement of Livestock, Animal and Poultry Science, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received: 28 September 1992 / Accepted: 11 February 1993

Abstract. Selection reduces additive genetic variation by generating gametic-phase disequilibrium, a phenomenon largely ignored when predicting response in plant breeding programs. The development of gametic-phase disequilibrium is here taken into account when predicting the response to selection for various schemes of recurrent selection applicable to plant populations. A general program permits prediction of response to selection from schemes of recurrent selection in which two or more rounds of selection occur in each cycle. An example from Sugar Beet, with alternate rounds of half-sib and S1 family selection, is illustrated. It is shown that failure to take into account the effects of gametic-phase disequilibrium can result in substantial overestimation of the response to selection as well as to changes in rank of the merits of alternative breeding schemes. For a given scheme, ignoring gametic-phase disequilibrium has only small effects on defining the optimum allocation of plots and the numbers of families tested.

Key words: Recurrent selection – Gametic phase disequilibrium – Sugar beet

Introduction

Recurrent selection in plants is defined as any scheme of cyclical selection by which the frequencies of favourable alleles are increased in plant populations (Hallauer 1981).

For any particular crop or population, the choice of a recurrent selection scheme depends on the breeder's objectives and such factors as cost, practical difficulties, and expected response to selection. It is clearly important then that the breeder be able to predict relative or absolute responses to selection from the various schemes he considers.

Given suitable estimates of genetic and environmental components of variation, formulae are available for the prediction of response to selection from many schemes (Sprague and Eberhart 1977; Hallauer 1981; Hallauer and Miranda 1981; Falconer 1989). However, such predicted responses are only valid for the first generation of selection, since selection itself will reduce additive genetic variation through the generation of gametic-phase disequilibrium (Falconer 1989), often referred to as the Bulmer effect (Bulmer 1971). Although this has been known for some time, and its consequences on the prediction of response to selection in animal breeding have been explored (e.g., Wray and Hill 1989), we are not aware of published work which discusses the effect of gametic-phase disequilibrium on recurrent selection schemes in plants.

The purpose of this paper is to extend the basic formulae for the prediction of response to selection to take linkage disequilibrium into account, and then demonstrate that it is important to consider the effects of linkage disequilibrium when choosing between alternative recurrent selection schemes.

For the sake of brevity, we shall not consider reciprocal recurrent selection schemes or schemes involving crossing to a tester. We shall also ignore complications due to genotype × environment interaction. The proposed theory can, however, be adapted to cover these complications. In addition, we shall consider only additive traits, since the consequences of linkage disequilibrium on non-additive variation are more complex. However, non-additive variation is not likely to change our

---

Communicated by A. R. Hallauer
Correspondence to: I. J. Mackay

conclusions, at least qualitatively (Falconer 1989; and see below).

Theory

A general procedure is developed for constructing deterministic simulations of plant breeding designs. The procedure identifies a series of recursive relationships which define the parental and Mendelian sampling contributions to genetic variation, the accuracy of selection, and from these the response to selection in each generation given the breeding and testing design.

Allowance is made for the reduction in genetic variance due to the selection of parents and the effect that this has on the accuracy of selection. Wherever possible, selection is based on information from collateral relatives, rather than ancestors or descendants. This allows the most parsimonious set of recursion relationships to be constructed.

The general form of the recursive relationships

In any generation the expected response to selection, $\varepsilon \mathrm{R}$, is given by

$$\varepsilon \mathrm{R} = 1/2\,(\mathrm{i}_1 \mathrm{r}_1 \sigma_{\mathrm{g}1} + \mathrm{i}_2 \mathrm{r}_2 \sigma_{\mathrm{g}2}) \tag{1}$$

where i_1 and i_2 are the intensities of selection in each sex of potential parents, r_1 and r_2 are the corresponding accuracies of selection, equal to the correlation between the true additive genetic value (breeding value) and the selection criterion, and $\sigma_{\mathrm{g}1}$ and $\sigma_{\mathrm{g}2}$ are the additive genetic standard deviations of each sex. In the case of equal selection intensity, criteria, and genetic variance in both sexes, equation 1 reduces to

$$\varepsilon \mathrm{R} = \mathrm{i\,r}\,\sigma_{\mathrm{g}}. \tag{1a}$$

The additive genetic value of an individual plant, g, is given by

$$\mathrm{g} = 1/2\,\mathrm{g}_{\mathrm{s}} + 1/2\,\mathrm{g}_{\mathrm{d}} + \mathrm{g}_{\mathrm{m}} \tag{2}$$

where g_{s} and g_{d} are the additive genetic values of the sire and dam, and g_{m} is the Mendelian sampling component.

From standard variance component theory, the expectation of the variance of g, $\varepsilon \sigma^2_{\mathrm{g}}$, is given by

$$\varepsilon \sigma^2_{\mathrm{g}} = 1/4\ \sigma^2_{\mathrm{g_s}} + 1/4\ \sigma^2_{\mathrm{g_d}} + \sigma^2_{\mathrm{g_m}} + 1/2\,\sigma_{\mathrm{g_s g_d}}. \tag{3}$$

With random mating of unrelated parental gametes the covariance term $\sigma_{\mathrm{g_s g_d}}$ is zero; $\sigma_{\mathrm{g_s g_d}}$ is non-zero in the presence of assortative or disassortative mating or matings of related plants, including selfing of parents. On the assumption of random selection of gametes from each parent, the expectation of the random Mendelian sampling component, $\varepsilon \sigma^2_{\mathrm{g_m}}$, is given by

$$\varepsilon \sigma^2_{\mathrm{g_m}} = 1/2\,(1 - 1/2\,\mathrm{F_s} - 1/2\,\mathrm{F_d})\,\sigma^2_{\mathrm{g_0}} \tag{4}$$

where $\mathrm{F_s}$ and $\mathrm{F_d}$ are the inbreeding coefficients of the sire and the dam, and $\sigma^2_{\mathrm{g_0}}$ is the additive genetic variance in the base population prior to any selection. As defined here, $\sigma^2_{\mathrm{g_s}}$ and $\sigma^2_{\mathrm{g_d}}$ refer to the genetic variances of the parents of the current generation. Thus, $\sigma^2_{\mathrm{g_s}}$ and $\sigma^2_{\mathrm{g_d}}$ are genetic variances after selection in the previous generation.

Following the theory of conditional variances as applied to genetic selection (Pearson, 1903), the expectation of genetic variance after selection, $\varepsilon \sigma^2_{\mathrm{g'}}$, is given by

$$\sigma^2_{\mathrm{g'}} = \sigma^2_{\mathrm{g}}\,(1 - \mathrm{k\,r}^2) \tag{5}$$

where r is the correlation between true additive genetic value and the selection criterion, referred to as the accuracy of selection, and $\mathrm{k} = \mathrm{i}\,(\mathrm{i} - \mathrm{x})$ where i is the standardized intensity of selection (Falconer, 1989) and x is the truncation point of the standardized normal distribution corresponding to a selection intensity i. Since this can be applied separately to each sex, $\sigma^2_{\mathrm{g_s}}$ and $\sigma^2_{\mathrm{g_d}}$ can be obtained from equation (5) given the intensity and accuracy of selection in each sex in the parental generation. Note that equation (5) is a more general form of that proposed by Cochran (1951) and Bulmer (1980), namely

$$\sigma^2_{\mathrm{g'}} = \sigma^2_{\mathrm{g}}\,(1 - \mathrm{k\,h}^2). \tag{5a}$$

Equation (5 a) applies only to selection on phenotype, or where σ^2_{g} and h^2 are re-defined to apply to group mean performance. However, equation (5) applies to any selection criterion and is therefore more widely applicable than (5 a) while also avoiding the necessity of redefining σ^2_{g} and h^2 to deal with selection on group means.

The accuracy of selection is given by

$$\mathrm{r} = \frac{\sigma_{\mathrm{g \cdot I}}}{\sqrt{\sigma^2_{\mathrm{g}}\,\sigma^2_{\mathrm{I}}}} \tag{6}$$

where $\sigma_{\mathrm{g \cdot I}}$ is the corvariance of an individual's genotype, g, with the selection criterion, I, and σ^2_{I} is the variance of the selection criterion. Extension to multitrait index selection, though not pursued here, is straightforward, since from standard theory of conditional covariances, the genetic covariance between two traits, $\sigma_{\mathrm{g_x g_y}}$, after selection on an index, I, is given by

$$\sigma'_{\mathrm{g_x g_y}} = \sigma_{\mathrm{g_x g_y}} - \frac{\sigma_{\mathrm{g_x I}}\,\sigma_{\mathrm{g_y I}}}{\sigma^2_{\mathrm{I}}} \cdot \mathrm{k}$$

General derivations for σ^2_{g} have been given, but derivations of $\sigma_{\mathrm{g \cdot I}}$ and σ^2_{I} depend on the mating scheme and the evaluation program envisaged and must be derived separately for each situation.

Equations (1) to (6) can be used to develop recursive prediction of the response to selection in each generation for most simple breeding designs, provided the information used in the selection decisions does not involve relatives more distant than half-sibs.

In order to extend these relationships to specific selection and mating designs, the phenotype of an individual plant in a testing program, P_i, is defined by

$$P_i = g_i + e1_i + e2_j \tag{7}$$

where $e1_i$ is an environmental effect unique to plant i in the given experimental plot j, and e 2 is an environmental effect common to all plants in plot j. The possibility of including non-additive genetic effects will be dealt with elsewhere.

Selection on the basis of half-sib mean performance

We illustrate the use of equations (1) to (6) by deriving expectations for response to selection and genetic variance for selection based on half-sib family means. No assumptions are made about the type of selection practised in the previous or following years. However, selection procedures and breeding structures in the previous generation determine the input values of $\sigma^2_{g_s}$, $\sigma^2_{g_d} \cdot F_s$ and F_d. It is assumed that the plant selected does not contribute to the mean performance of the half-sib family mean. Each plot in the half-sib group is assumed to have a separate unidentified male parent and there are m plants per plot and n plots per family. Assuming random mating of parents, the expectation of genetic variance in this generation is, from equations (3) and (4), given by

$$\varepsilon\sigma^2_g = 1/4\,\sigma^2_{g_s} + 1/4\,\sigma^2_{g_d} + 1/2\,(1 - 1/2\,F_s - 1/2\,F_d)\,\sigma^2_{g_0}. \tag{8}$$

The half-sib family mean performance, $P_{\overline{HS}}$, is

$$P_{\overline{HS}} = 1/2\,g_d + \frac{\sum^{mn}(1/2\,g_{s_i} + g_{m_i} + e1_i)}{mn} + \frac{\sum^{n} e2_j}{n} \tag{9}$$

and the expectation of its variance is

$$\varepsilon\sigma^2_{P_{\overline{HS}}} = 1/4\,\sigma^2_{g_d} + \frac{1/4\,\sigma^2_{g_s} + \sigma^2_{g_m} + \sigma^2_{e1}}{mn} + \frac{\sigma^2_{e2}}{n} \tag{10}$$

The covariance between the genetic value of an individual, g_i, and the selection criterion, $P_{\overline{HS}}$, with expectations as defined in equations (2) and (7), assuming random mating of parents, is therefore given by

$$\varepsilon\sigma_{g\cdot P_{\overline{HS}}} = 1/4\,\sigma^2_{g_d}. \tag{11}$$

Following from equation (6), the accuracy of selection is given by

$$r = \frac{\sigma_{g\cdot P_{\overline{HS}}}}{\sqrt{\sigma^2_g\,\sigma^2_{P_{\overline{HS}}}}}$$

which by substituting from equations [8], [10] and [11], gives

$$\varepsilon r = \frac{\sigma^2_{g_d}}{4\sqrt{\sigma^2_g\left[1/4\,\sigma^2_{g_d} + \dfrac{1/4\,\sigma^2_{g_s} + \sigma^2_{g_m} + \sigma^2_{e1}}{mn} + \dfrac{\sigma^2_{e2}}{n}\right]}} \tag{12}$$

If both sexes are treated equally, equations (8) and (12) reduce to

$$\varepsilon\sigma^2_g = 1/2\,\sigma^2_{g_p} + 1/2\,(1 - F_p)\,\sigma^2_{g_0} \tag{8a}$$

and

$$\varepsilon r = \frac{\sigma^2_{g_p}}{4\sqrt{\sigma^2_g\left[\dfrac{\sigma^2_{g_p}(mn+1) + 4(\sigma^2_{e1} + \sigma^2_{g_m} + m\,\sigma^2_{e2})}{4\,mn}\right]}} \tag{12a}$$

Response to selection and genetic variance after selection (equal to the genetic variance of parents of the following generation) follow directly from equations (1) or (1a) and (5).

Provided that population sizes are sufficiently large, the inbreeding coefficient of parents of the following generation are approximately

$$F_{s_{t+1}} = F_{d_{t+1}} = 0.$$

Selection for other types of family structure

Derivations of formulae for response to selection and genetic variance for other forms of breeding and selection follow the pattern laid out for half-sib selection above. Common forms of plant selection would be full-sib family selection, S1 family selection (progeny of selfed parents), clonal selection, doubled haploid selection, and mass selection. A point to note with doubled haploids is that parental gametes are not selected at random so that equation (4) does not hold and must be replaced by

$$\varepsilon\sigma^2_{g_m} = (1 - F_p)\,\sigma^2_{g_0}. \tag{4a}$$

The simulation program

A deterministic simulation program was constructed in SAS IML (SAS, 1988). The program contains modules corresponding to half-sib (HS), full-sib (FS), S1 and clonal family mean selection, random mating (no selection), and the generation of doubled haploids (DH) (no selection). The program accepts as input parameters, definitions of the types of generations and selection criteria (see below) and the order in which they appear in a repeated selection cycle, the number of repetitions of the cycle, σ^2_{e1}, σ^2_{e2} and σ^2_{g0}, and the number of plants per plot, m, and plots per family, n, for each type of generation. Mass selection can be achieved by setting m and n to 1 in clonal selection. Doubled haploid selection can be achieved by generating doubled haploids followed by a round of clonal or S1 family selection.

The number of families tested, n_{ft}, and selected, n_{fs}, are entered and from these the selection intensity is derived using Burrows (1972) approximation for selection intensities in finite samples as

$$i = i_\infty - (1 - p)/2\,i_\infty\,p\,(n_{ft} - 1)$$

where i_∞ is the standardized selection differential assuming infinite population size, and p is the proportion selected. Note that this formulation assumes that selection applies across families with an equal number of individuals selected per family. This is appropriate with the examples examined here, but would not apply if more complex selection criteria, such as combined full and half-sib indexes, were employed.

The program starts with a base population with no selection history and genetic variances in gametic-phase equilibrium. Successive generations are chosen according to the order defined in the inputs. Each generation defines selection response in that generation (if selection is practised) and the genetic variances and inbreeding coefficient of the parents of the following generation.

Examples

One round of selection per cycle

The formulae described above can be applied to many of the schemes of selection commonly encountered in plant breeding. Some examples are given in Table 1 where responses to selection from various schemes of family selection in a previously unselected population are listed. Variance components ($\sigma^2_{g0} = 0.5$, $\sigma^2_{e1} = 12.25$, $\sigma^2_{e2} = 0.25$), the number of plants per plot (m = 50), the number of replicate plots (n = 4), and the proportion selected (10%) have been chosen to approximate estimates of variance components for root yield in a population of Sugar Beet maintained by the author (I. J. Mackay). However, these parameters are also suitable for illustrative purposes since they are within the ranges frequently found or used in practice for many crops. Response to selection in the first cycle of selection is given in the first column of Table 1. These responses correspond to those given by the classical formulae of Hallauer (1981) and others.

Column 2 lists responses to selection at equilibrium, where response to selection per cycle is constant and the loss of variation due to linkage disequilibrium is balanced by the generation of variation through segregation. Equilibrium is attained within five generations, with responses to selection changing very little after the first two cycles of selection.

Loss of variation due to gametic-phase disequilibrium is temporary. One half of the reduction in variance is restored following each round of random mating in the absence of selection. Schemes marked with a * in Table 1 are schemes in which each generation of selection followed by random mating is followed by a second generation of random mating without selection. Such schemes have been proposed to reduce the effect of physical linkage on selection response (reviewed by Hallauer and Miranda 1981), whereas here we assume no physical linkage of loci controlling genetic variance (i.e., the infinitesimal model).

Table 1. Response to selection in the first cycle and at equilibrium for various schemes of recurrent selection[a]

Scheme	R1	Req	σ^2_{gf1}	σ^2_{gfeq}
Half-sib families	0.436	0.384	0.125	0.105
Half-sibs*	0.436	0.419	0.125	0.117
Full-sib families	0.714	0.547	0.250	0.170
Full-sibs*	0.714	0.642	0.250	0.213
S1 progenies	1.107	0.831	0.500	0.315
S1 progenies*	1.107	0.986	0.500	0.413
Doubled haploid progenies	1.652	1.229	1.000	0.595
Doubled haploids*	1.652	1.466	1.000	0.807

[a] $\sigma^2_g=0.5$, $\sigma^2_{e1}=12.25$, $\sigma^2_{e2}=0.25$, m=50, n=4 for all schemes

* Denotes schemes in which each generation of selection and random mating is followed by a second generation of random mating without selection, R1 = response to selection in the first cycle, Req = response to selection at equilibrium, σ^2_{gf1} is the between-family component of genetic variance before selection in the first cycle, σ^2_{gfeq} is the between-family component of genetic variance before selection at equilibrium. See text for details

For the doubled haploid schemes, it is assumed that doubled haploids are derived from F_1 crosses produced by random mating selected doubled haploids from the previous cycle or from four-way crosses produced by random mating of the F_1s in the case of the scheme marked "*"

Responses to half-sib selection are readily calculated from the second section on Theory. The calculation of response to full-sib selection is also simply derived from the first section on Theory. The other schemes listed in Table 1 all involve more than one generation per cycle. Response to selection from these schemes involves successive application of formulae for the family types being generated. These formulae are readily derived from the first section on Theory.

Columns 3 and 4 of Table 1 give the between-family component of additive genetic variance in the selection generation for the various schemes, illustrating the magnitude of the reduction in genetic variation caused by gametic-phase disequilibrium. Although not shown, this reduction is achieved very quickly, being very close to its equilibrium value after two cycles of selection.

Several points are clear from Table 1. Firstly, the ranking of schemes for response to selection per cycle is the same in the first as in later generations; DH > S1 > FS > HS. However, the magnitudes of the differences between the schemes are reduced substantially. For example, FS : HS response in generation 1 is 1.63, but at equilibrium is 1.42. Thus, with optimization of plot size and number, and taking into account differences in costs and cycle times, economic ranking of different schemes could depend on whether response was estimated by ignoring or including gametic-phase disequilibrium.

As expected, response per cycle increases following a round of random mating without selection. This is most extreme for the doubled haploid and S1 progeny scheme. Often the extra round of random mating is unlikely to be worthwhile, since in many crops this extra generation will involve an extra year. Thus, although response per cycle will rise, the response per year will usually fall. Doubled haploids may be an exception. Here, response per cycle is increased by about 20% by allowing an additional generation of random mating. Starting with an F1 cross, it will often take about 5 years to cycle through doubled haploid production, multiplication, yield trials, selection, and the production of second cycle F1s by crossing selected doubled haploids. Thus, even if an extra round of random mating involves a full year, response to selection per year remains constant while the period between expensive yield trials increases from 5 to 6 years. In crops without a vernalisation requirement, an additional round of random mating may be possible without extension of the cycle time and with little additional cost.

Table 2. Maximum response[a] per year with S1-HS selection cycle, accounting for or not accounting for gametic-phase disequilibrium when $V_a = 0.5$[b]

# plot S1	# rep S1	With gpd[c]		Without gpd	
		R max	# rep[d] HS	R max	# rep HS
300	1	0.305	2	0.428	3
	2	0.269	2	0.388	3
	3	0.235	3	0.341	3
	4	0.205	3	0.292	3
	5	0.177	3	0.243	3
400	1	0.318	2	0.445	2
	2	0.286	2	0.415	2
	3	0.257	2	0.377	2
	4	0.232	2	0.339	2
	5	0.208	2	0.301	2
500	1	0.324	2	0.455	2
	2	0.298	2	0.432	2
	3	0.272	2	0.399	2
	4	0.249	2	0.366	2
	5	0.227	2	0.334	2
600	1	0.328	1	0.459*	2
	2	0.305	2	0.441	2
	3	0.287	2	0.412	2
	4	0.254	2	0.382	2
	5	0.240	2	0.354	2
700	1	0.330*	1	0.456	2
	2	0.307	1	0.442	2
	3	0.285	2	0.416	2
	4	0.265	2	0.389	2
	5	0.246	2	0.363	2
800	1	0.328	1	0.449	1
	2	0.308	1	0.438	1
	3	0.287	1	0.414	1
	4	0.267	1	0.389	1
	5	0.250	1	0.365	1

[a] Maximum response at a given number of plots allocated to S1 testing and a given number of replicate plots per S1 family

[b] Other parameters: 1,000 plots per cycle, $V_{e1} = 12.5$, $V_{e2} = 0.25$, 50 plants per plot, 50 S1 and 25 HS families selected

[c] gpd = gametic-phase disequilibrium

[d] As the number of replicates per S1 and HS family increases, the number of families tested decreases to match total number of plots available

* Indicates overall maximum response

Two or more rounds of selection per cycle

Extension of the method to deal with multiple generations of selection is illustrated with an example from Sugar Beet where S1 selection can be followed either by a generation of half-sib selection (S1-HS scheme) or a generation of random mating (S1-only scheme). The reasons for using these schemes are outlined in Appendix 1.

Bradshaw (1984) presented computer simulations of responses to selection from this and similar schemes. Dhillon (1991) published formulae for the response to selection from this scheme which failed to take into account either the reduction in variation as a result of selection or the increase in response to half-sib selection through using inbred parents. More accurate predictors of response for an additive character can be obtained by applying the already given formulae for half-sib selection, alternately with similarly derived formulae for S1 family selection as illustrated below.

Environmental variances and physical parameters are assumed to be the same as used in Table 1. Three levels of initial genetic variances are assumed; $V_a = 0.1$, 0.5 or 1.0. The value $V_a = 0.5$ has been found by personal experience (I. J. Mackay) to represent a commercial Sugar Beet population. The other two values have been included to study the sensitivity of our conclusions to changes in genetic variance. The total plots were constrained to 1,000 per cycle, where a cycle consists of one generation of S1 selection followed by one generation of half-sib selection. There were 50 S1 progenies selected which generated half-sib families and 25 half-sib families selected which generated the following S1 progenies. Again, these parameters approximate commercial Sugar Beet breeding. Equal numbers of individuals are taken from each family selected. This scheme should keep the effects of drift and inbreeding to acceptable levels, at least in the short and medium term. Of interest here are the effects of including gametic-phase disequilibrium on the optimum partition of a fixed test resource between generations and the optimum number of families and replicates tested per family, as well as on the estimated gain in selection response through including selection of half-sib families compared to selection between S1 families alone. The number of plots per generation (S1 vs HS) and the number of replicate plots per family were optimized to obtain maximum response to selection at equilibrium by varying the number of S1 plots from 300 to 800 and the number of replicate plots for both S1 and HS families independently from 1 to 5. With S1 selection only, 1,000 plots were tested every cycle and 50 S1 families selected.

Table 2 shows the predicted maximum response, and the corresponding number of replicate HS plots that achieve maximum response, for varying combinations of the number of S1 plots and the number of replicate plots per S1 family, where gametic-phase disequilibrium is or is not taken into account, when the initial genetic variance ($\sigma^2_{g_0}$) is 0.5. Similar calculations were made when $\sigma^2_{g_0} = 0.1$ and 1.0 (data not shown). Maximum responses occurred with 600 or 700 S1 plots (i.e., 400 or 300 HS plots) and either one or two replicate plots per S1 family and 1–3 replicate plots per HS family. Although not shown in Table 2, the response was not sensitive to changing the number of replicate plots per HS family, varying by less than 1.5% as the number of replicates was altered by ±1 from the optimum.

Table 3. Equilibrium response per year with an S1-only selection cycle, with gametic-phase disequilibrium accounted for and not accounted for when $V_a = 0.1$, 0.5 or 1.0[a]

V_a	# rep S1	Equilibrium response rate	
		With gpd[b]	Without gpd
0.1	1	0.079	0.088
	2	0.084	0.099
	3	0.082	0.100
	4	0.080	0.098
	5	0.076	0.094
0.5	1	0.270	0.342
	2	0.257	0.336
	3	0.240	0.316
	4	0.223	0.294
	5	0.208	0.272
1.0	1	0.424	0.558
	2	0.391	0.521
	3	0.359	0.477
	4	0.330	0.438
	5	0.304	0.402

[a] Other parameters: 1,000 plots per cycle, $V_{e1} = 12.5$, $V_{e2} = 0.25$, 50 plants per family, 50 S1 families selected

[b] gpd = gametic-phase disequilibrium

Table 3 shows the predicted response when selection is on S1 families with an intervening generation of random selection and mating (S1-only scheme), when gametic-phase disequilibrium is and is not accounted for. In this case, accounting for gametic-phase disequilibrium had a small effect on the optimum design when $V_a = 0.1$ but no effect for other values of V_a. As expected, absolute rates of response are predicted to be lower when gametic-phase disequilibrium is accounted for.

Table 4 shows the optimum design and maximum response rate for different schemes (S1-Hs vs S1-only) at the three levels of genetic variance, with and without accounting for gametic-phase disequilibrium. Also shown are the ratios of maximum predicted annual response for the S1-HS scheme compared to the S1-only scheme, along with the ratios of predicted response for models with and without gametic-phase disequilibrium.

The S1-HS scheme always out-performs the S1-only scheme. However, failure to account for gametic-phase disequilibrium leads to substantial over-prediction of the advantage of S1-HS schemes over S1-only schemes, par-

Table 4. Maximum equilibrium response per year with S1 or S1-HS selection cycles, and conditions leading to that response for $V_a = 0.1$, 0.5 or 1.0 with models accounting for and not accounting for gametic-phase disequilibrium

Selection cycle	V_a	# plot S1	# rep S1	# rep HS	R[a]	Response ratio	
						S1-HS/S1	Zgpd/gpd[b]
S1-HS selection							
Bulmer effect present	0.1	700	2	2	0.100	1.19	–
	0.5	700	1	1	0.330	1.22	–
	1.0	700	1	1	0.520	1.23	–
No Bulmer effect	0.1	600	2	3	0.123	1.23	1.23
	0.5	600	1	2	0.459	1.34	1.39
	1.0	700	1	1	0.759	1.36	1.46
S1 selection only							
Bulmer effect present	0.1	1,000	2	0	0.084	–	–
	0.5	1,000	1	0	0.270	–	–
	1.0	1,000	1	0	0.424	–	–
No Bulmer effect	0.1	1,000	3	0	0.100	–	1.19
	0.5	1,000	1	0	0.342	–	1.27
	1.0	1,000	1	0	0.558	–	1.31

[a] R = equilibrium response per year

[b] Zgpd/gpd = response without gametic-phase disequilibrium divided by response with gametic-phase disequilibrium accounted for

Table 5. Response per cycle and genetic variance in each generation at cycles 1, 2 and 5 of selection with S1-HS or S1-only selection schemes when gametic-phase disequilibrium is accounted for and not accounted for, and each scheme optimized as in Table 4

Selection scheme	Cycle 1			Cycle 2			Cycle 5		
	Resp.[a]	σ^2_{g1}[b]	σ^2_{g2}[c]	Resp.[a]	σ^2_{g1}	σ^2_{g2}	Resp.[a]	σ^2_{g1}	σ^2_{g2}
With gametic-phase disequilibrium									
$V_a=0.1$									
S1-HS	0.355	0.150	0.089	0.316	0.136	0.084	0.301	0.130	0.082
S1 only[d]	0.296	0.150	0.088	0.265	0.138	0.084	0.252	0.133	0.083
$V_a=0.5$									
S1-HS	1.259	0.750	0.396	1.040	0.622	0.370	0.990	0.595	0.364
S1 only	1.026	0.750	0.394	0.856	0.644	0.373	0.811	0.618	0.368
$V_a=1.0$									
S1-HS	2.050	1.500	0.723	1.621	1.162	0.675	1.559	1.117	0.667
S1 only	1.675	1.500	0.718	1.333	1.218	0.681	1.275	1.174	0.673
Without gametic-phase disequilibrium									
$V_a=0.1$									
S1-HS	0.370	0.150	0.100						
S1 only[d]	0.300	0.150	0.100						
$V_a=0.5$									
S1-HS	1.377	0.750	0.500	Values in other cycles identical to cycle 1					
S1 only	1.026	0.750	0.500						
$V_a=1.0$									
S1-HS	2.278	1.50	1.00						
S1 only	1.675	1.50	1.00						

[a] Response per cycle

[b] Genetic variance at S1 selection

[c] Genetic variance at HS selection (S1-HS scheme) or random selection (S1-only scheme)

[d] Responses for 1st cycle S1-only with and without gametic-phase disequilibrium are not the same due to small differences in optimum designs for equilibrium response rates

ticularly at high levels of genetic variance. Failure to account for gametic-phase disequilibrium leads to over prediction of S1 responses by 19 to 31% depending on genetic variance, and over-prediction of S1-HS responses by 23 to 46%.

The reason for these differences in levels of over-prediction can be seen in Table 5, where responses and levels of genetic variation in cycles 1, 2 and 5 of selection are shown. Genetic variance is reduced by selection to between 86 and 67% of the original levels depending on the initial levels of genetic variation and the generation in question (S1 or intervening generation).

Genetic variance is reduced most in both absolute and proportional terms at high levels of genetic variation because of the higher accuracy of selection. Hence, biases in estimated response rates when not accounting for gametic-phase disequilibrium are greatest at high initial levels of genetic variance. Of particular interest here is the ratio of genetic variance in the S1 generation to that in the intervening generation (random selection or HS selection). Ignoring the effect of gametic-phase disequilibrium, this ratio is predicted to be 1.5 but, in the presence of disequilibrium, this ratio is between 1.58 and 1.74 depending on the initial genetic variance and the type of selection scheme. Thus ignoring gametic-phase disequilibrium leads to overestimation of the relative amount of genetic variance available for selection in the intervening (HS) generation and hence over-prediction of the relative merit of S1-HS schemes versus S1-only schemes.

Discussion

The theory developed above improves the precision with which response to selection can be predicted, and permits more accurate consideration of alternative breeding schemes. In particular, for the first time it has been possible to predict responses to selection from alternative generations of half-sib and S1 family selection. The formulae for these schemes are not complex and can easily be built into a computer simulation program or a computer spreadsheet.

Although our examples consider only single trait selection and between family selection, extension to multiple trait index selection is straightforward. Similarly, the theory permits the prediction of response to within-family selection and consequently the development of selection indices combining information from relatives in the prediction of individual genetic merit. Expectations for within-family variances are easily derived for any of the schemes discussed above by subtraction of the between-family genetic component of variation from the total genetic variance. Although most plant breeding programmes do not make use of information from relatives, there are important exceptions. For example, Gallais (1989) has considered the optimization of recurrent selection programmes based on selection between doubled haploids using an index incorporating information from doubled haploids derived from the same parent. Our method can readily be used to assess the effects of gametic-phase disequilibrium on such schemes. Similarly, with S1 family selection, between-family variation can be partitioned into between half-sib or full-sib pedigrees and between S1 families within pedigrees. There is, therefore, the opportunity to select between S1 families on an index which includes information from relatives.

A limitation of the present development is that it fails to accommodate non-additive variation. However, many of the schemes of selection routinely used by plant breeders are little affected by dominance. Of those considered here, only for those involving selection between full-sib families or S1 families is there a need to consider the consequences of dominance variation. For full-sib families, the dominance component is relatively small, and it seems unlikely that failure to consider it will seriously bias results (Falconer 1989). For S1 family selection, the dominance component of variation is complicated (Empig et al. 1972; Jan-orn et al. 1976; Eckebil et al. 1977; Bradshaw 1983). The consequences of non-additive variation on the generation of gametic-phase disequilibrium in recurrent selection programmes involving selfing is an important area for further study.

The effects of drift on responses to selection are also ignored here. This is indeed a limitation, since different breeding schemes will have different effective population sizes (Sprague and Eberhart 1977; Choo and Kannenberg 1979; Hallauer and Miranda 1981). However, in the short to medium term, the reduction in variance due to gametic-phase disequilibrium is of much greater importance. A general theory taking into account the effects of both drift and gametic-phase disequilibrium will be difficult, although a start has been made in animal breeding, as reviewed by Verrier et al. (1991).

The reductions in selection response of 16–32% due to gametic-phase disequilibrium found here are larger than the 15–25% typically found for animal breeding schemes (Keller et al. 1989; Meyer and Smith 1990; Dekkers 1992). In contrast to these results, gametic-phase disequilibrium has relatively little effect on the difference in response for alternative animal breeding schemes, as seen in the published literature (Keller et al. 1989; Meyer and Smith 1990) and from personal experience (J. P. Gibson). Thus, accounting for gametic-phase disequilibrium, although accepted as an essential part of modelling in animal breeding, may well be of more practical importance in plant breeding.

Appendix

In Sugar Beet, by establishing populations fixed for an obligate self-fertility gene, but segregating for a recessive male-sterility gene, it is possible to carry out recurrent selection schemes based on S1 progeny testing (Bosemark 1971). To generate selfed progenies, one marks and harvests the fertile plants. Recombination of selected selfed progenies is possible by mixing plants of selected lines and harvesting seed from the sterile segregants. Similar schemes have been proposed for many crops, reviewed by Rao et al. (1990), and are used by a number of Sugar Beet breeders (Bosemark 1971; Lewellen 1984; Lewellen and Skoyen 1989). In Sugar Beet this scheme takes 3 years per cycle: a year to self progenies, a year for yield trials and a year for recombination of selected progenies. However, it is possible, with no loss of time, to extend the scheme by including a round of half-sib selection. For this, the half-sib seed of male-sterile plants harvested in the first year is tested in trials in the second year. Also, in the second year, seed is produced from obligate self-fertile individuals in each half-sib family. In the third year, S1 progenies from selected half-sib families are tested in yield trials. The fourth year is then equivalent to the first year of the following cycle in which individuals from selected S1 progenies are recombined and half-sib families for testing in the following year are generated. Although no increase in cycle time is involved by including selection between half-sib families, there is considerable extra expense since with a proportion p of half-sib families being selected, 1/p times more S1 progenies are produced than are required for testing.

References

Bosemark NO (1971) Use of Mendelian male-sterility in recurrent selection and hybrid breeding in beets. Eucarpia Fodder Crops Section. Report of meeting at Lusignan, France, pp 127–135

Bradshaw JE (1983) Estimating the predicted response to S1 family selection. Heredity 51:415–418

Bradshaw JE (1984) Computer simulation of family selection suitable for Kale (*Brassica oleracea* L.), involving half-sib, full-sib and selfed families. Theor Appl Genet. 68:503–508

Burrows PM (1972) Expected selection differentials for directional selection. Biometrics 28:1091–1100

Bulmer MG (1971) The effect of selection on genetic variability. Am Nat 105:201–211

Bulmer MG (1980) The mathematical theory of quantitative genetics. Oxford University Press, Oxford

Choo TM, Kannenberg LW (1979) Changes in gene frequency during mass, modified ear-to-row, and S1 selection: a simulation study. Crop Sci 19:503–509

Cochran WG (1951) Improvement by means of selection. In: Neyman J (ed) Proc 2nd Berkeley Symp Math Stat Prob University of California Press, Berkeley, pp 449–470

Dekkers JCM (1992) Asymptotic response to selection on best linear unbiased predictors of breeding values. Anim Prod 54:351–360

Dhillon BS (1991) Alternate recurrent selection of S1 and half-sib families for intrapopulation improvement. Maydica 36:45–48

Eckebil JP, Ross WM, Gardener CO, Maranville JW (1977) Heritability estimates, genetic correlations, and predicted gains from S1 progeny tests in three grain sorghum random-mating populations. Crop Sci 17:373–377

Empig LT, Gardener CO, Compton WA (1972) Theoretical gains for different population improvement procedures. Nebraska Agricultural Experiment Station Bulletin 26 (revised)

Falconer DS (1989) Introduction to quantitative genetics (3rd edn.) Longman, Harlow, England

Gallais A (1989) Optimization of recurrent selection on the phenotypic value of doubled haploid lines. Theor Appl Genet 77:501–504

Hallauer AR (1981) Selection and breeding methods. In: Frey KJ (ed) Plant breeding II. Iowa State University, Ames, Iowa, pp 3–55

Hallauer AR, Miranda JB (1981) Quantitative genetics in maize breeding. Iowa State University Press, Ames, Iowa

Jan-orn J, Gardener CO, Ross WM (1976) Quantitative genetic studies of the NP3R random-mating grain sorghum population. Crop Sci 16:489–496

Keller DS, Gearheart WW, Smith C (1989) A comparison of factors reducing selection response in closed nucleus breeding schemes. J Anim Sci 68:1553–1561

Lewellen RT (1984) Breeding sugarbeet for crop efficiences. Proc Calif Plant and Soil Conf, p 165

Lewellen RT, Skoyen IO (1989) S1 progeny recurrent selection to improve yield in sugarbeet. J Sugar Beet Res 26:A14

Meyer K, Smith C (1990) Comparison of theoretical and simulated equilibrium genetic response rates with progeny testing in dairy cattle. Anim Prod 50:207–212

Pearson K (1903) Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. Philos Trans R Soc Lond Ser A 200:1–66

Rao MK, Devi KU, Arundhati A (1990) Applications of genetic male sterility in plant breeding. Plant Breed 105:1–25

Sprague GF, Eberhart SA (1977) Corn breeding. In: Sprague GF (ed) Corn and corn improvement. Am Soc Agron, Madison, Wisconsin, Chapter 6, pp 305–362

Verrier E, Colleau JJ, Foulley JL (1991) Methods for predicting response to selection in small populations under additive genetic models: a review. Livestock Prod Sci 29:93–114

Wray NR, Hill WG (1989) Asymptotic rates of response from index selection. Anim Prod 49:217–227