Theor Appl Genet (1986) 73: 182–189

Selection theory for selfed progenies*

N. M. Cowen**

Department of Agronomy, University of Missouri, Columbia, MO 65211, USA

Received November 11, 1985; Accepted August 19, 1986
Communicated by A. R. Hallauer

Summary. The purpose of this article was to extend the model used to predict selection response with selfed progeny from 2 alleles per locus to a model which is general for number and frequency of alleles at loci. To accomplish this, 4 areas had to be dealt with: 1) simplification of the derivation and calculation of the condensed coefficients of identity; 2) presentation of the genetic variances expressed among and within selfed progenies as linear function of 5 population parameters; 3) presentation of selection response equations for selfed progenies as functions of these 5 population parameters; and 4) to identify a set of progeny to evaluate, such that one might be able to estimate these 5 population parameters.

The five population parameters used in predicting gains were the additive genetic variance, the dominance variance, the covariance of additive and homozygous dominance deviations, the variance of the homozygous dominance deviations and a squared inbreeding depression term.

Key words: Selection response – Selfed progeny – Variance – Dominance deviations – Inbreed depression

Introduction

Several models have been developed which explain genotypic covariances of relatives under some form of inbreeding (Gillois 1964; Harris 1964; Jacquard 1974; Weir and Cockerham 1977). These models are applicable to random mating populations, with multiple alleles per locus, and are general for system of inbreeding. Several models have also been developed for the more restricted inbreeding system of self-fertilization (Horner and Weber 1956; Cockerham 1983). Horner and Weber's model requires the assumption of two alleles per locus. Each of these models can be used to determine the expectations of genotypic covariances of inbred individuals with known relationship. Cockerham (1983) has shown how his model may be used to determine the total genetic variance expressed among individuals at a given level of inbreeding and the genotypic covariance of parent and offspring. In contrast, crop breeders work primarily with progenies obtained by self-fertilization, and are interested in genetic variability expressed among and within these progenies.

Theoretical gains from selection can be calculated as the product of the selection differential and the regression of the phenotypic value of the response unit on the phenotypic value of the selection unit (Hallauer and Miranda 1981). The theoretical models used in prediction, where both the response units and selection units are non-inbred progenies or individuals, are general for numbers and frequencies of alleles at loci (Cockerham 1969; Nyquist 1978). This, however, is not the situation when both the response and selection units are inbred progenies. The prediction equations for either S_1 or S_2 recurrent selection are based on models which assume two alleles per locus (Sprague 1966).

Recently, Cockerham and Matzinger (1985) presented selection response equations, based on a model which is general for number and frequency of alleles, where the selection unit was an inbred individual and the response unit was either an individual who is more inbred or an outbred individual. In contrast, the equations presented by Sprague (1966) predict the change in the population mean following both recombination and inbreeding to an identical level as the

---

* Contribution from the Missouri Agricultural Experiment Station. Journal Series No. 9971

** Current address: United Agriseeds, P.O. Box 4011, Champaign, Ill 61820, USA selection unit. The only available information on estimating the parameters used by Cockerham (1983) was a suggested set of progenies to evaluate.

The purpose of this article is to present an alternate derivation of the identity measures employed by Cockerham (1983) and Jacquard (1974); present the expectations of variances expressed among and within selfed progenies; present selection response equations where the selection and response units are equally inbred; and to provide an alternative set of progeny to evaluate which can provide estimates of the parameters in Cockerham's model (Cockerham 1983).

Alternate derivation of condensed coefficients of identity

The covariance of relatives under any regular system of inbreeding can be calculated as a linear function of five population variances or quadratic forms (Cockerham 1983; Jacquard 1974). The coefficients of each population parameter are either sums or differences of the condensed coefficients of identity defined by both Cockerham (1983) and Jacquard (1974). The condensed coefficients of identity are tedious and often difficult to calculate directly. Nyquist (1978) presented a set of 37 rules used in the direct calculation of Cockerham's condensed coefficients of identity. Indirect methods of calculation have been proposed by Chevalet and Gillois (1977) and used by Cowen et al. (1985).

When the inbreeding system is strictly limited to self pollination, then the condensed coefficients of identity can be calculated directly as products of the inbreeding coefficient and conditional inbreeding coefficients. This alternative derivation is presented in Table 1.

It should be noted that condensed states of identity 6 through 9 cannot occur with strict self pollination for individuals in generations g and g′ where the last common ancestor is in generation t. All 4 of these states of identity require a minimum of 3 alleles all mutually nonidentical by descent, however the individual in generation t carries only two alleles. Calculation of the inbreeding coefficients of the individuals in generations g and g′, the coefficient of parentage between them, and the coefficients of all the population parameters in the genotypic covariance of these two relatives can now be accomplished. The inbreeding coefficients are obtained as either

$$F_g = 1 - (\tfrac{1}{2})^g \quad \text{and}$$

$$F_{g'} = 1 - (\tfrac{1}{2})^{g'}, \quad \text{or}$$

$$F_g = \Delta_1 + \Delta_2 + \Delta_4 \quad \text{and}$$

$$F_{g'} = \Delta_1 + \Delta_3 + \Delta_4 .$$

The coefficient of parentage between these individuals in generations g and g′ is

$$r_{gg'} = \Delta_1 + \tfrac{1}{2}(\Delta_2 + \Delta_3 + \Delta_5) = \tfrac{1}{2}(1 + F_t) .$$

Where $r_{gg'}$ is designated $\theta_{gg'}$ by Cockerham (1983). Before the covariance of these individuals can be defined the population variances and quadratic components involved must first be defined.

The population considered is assumed to be in linkage equilibrium, with no linkage of genes affecting the trait. Assuming no epistasis, then, all variances and quadratic components defined for the population will be the sum of the values for the individual loci, and consideration can be given to a single locus. Similarly, predicted selection response can be expressed as the sum of the predicted responses for individual loci. The structure of the population at a single locus is

$$\sum_i \sum_j p_i p_j A_i A_j ,$$

and the structure after n generations of self pollination is

$$F_n \sum_i p_i A_i A_i + (1 - F_n) \sum_i \sum_j p_i p_j A_i A_j$$

where p_i is the frequency of A_i in the population and F_n is defined as above. The structure of the fully inbred population is $\sum_i p_i A_i A_i$. Assuming an additive, dominance model, the genotypic value for $A_i A_j$ is

$$G_{ij} = \mu + \alpha_i + \alpha_j + \delta_{ij} ,$$

where μ is the population mean genotypic value, α_i and α_j are the average effects of alleles A_i and A_j fit by least squares procedures, and δ_{ij} is the dominance deviation. In the remaining discussion, the genotypic values will be coded by subtracting μ. By definition,

$$\sum_i p_i \alpha_i = \sum_j p_j \alpha_j = 0 , \quad \text{and}$$

$$\sum_i p_i \delta_{ij} = \sum_j p_j \delta_{ij} = \sum_i \sum_j p_i p_j \delta_{ij} = 0 .$$

It can be shown that the genotypic covariance of relatives is a linear function of five population variances or quadratic components; the additive genetic variance

$$\sigma_A^2 = 2 \sum_i p_i \alpha_i^2 ,$$

the dominance variance

$$\sigma_D^2 = \sum_i \sum_j p_i p_j \delta_{ij}^2 ,$$

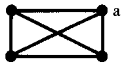
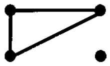
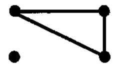
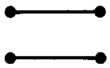
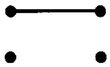
the covariance of additive and homozygous dominance deviations

$$D_1 = \sum_i p_i \alpha_i \delta_{ii} ,$$

the variance of the homozygous dominance deviations

$$D_2 = \sum_i p_i \delta_{ii}^2 - \Big[\sum_i p_i \delta_{ii}\Big]^2 ,$$

Table 1. Condensed coefficients of identity for individuals in generations g and g′ obtained by self pollination where the last common relative was in generation t

State no.	State of identify of progeny in generations	Identity condition of parent in generation t			
	g g′	● ●	●	●	Condensed coefficient
1.	 [a]	$\frac{1}{2}(F_{g-t})(F_{g'-t})$	1	Δ_1	$(1-F_t)\frac{1}{2}(F_{g-t})(F_{g'-t})+F_t$
2.		$(F_{g-t})(1-F_{g'-t})$	0	Δ_2	$(1-F_t)(F_{g-t})(1-F_{g'-t})$
3.		$(1-F_{g-t})(F_{g'-t})$	0	Δ_3	$(1-F_t)(1-F_{g-t})(F_{g'-t})$
4.		$\frac{1}{2}(F_{g-t})(F_{g'-t})$	0	Δ_4	$(1-F_t)(\frac{1}{2})(F_{g-t})(F_{g'-t})$
5.		$(1-F_{g-t})(1-F_{g'-t})$	0	Δ_5	$(1-F_t)(1-F_{g-t})(1-F_{g'-t})$
6.		0	0	Δ_6	0
7.		0	0	Δ_7	0
8.		0	0	Δ_8	0
9.		0	0	Δ_9	0
Probability of parental state		$(1-F_t)$	(F_t) [b]		

---

[a] A line connecting any pair of alleles indicates they are identical by descent

[b] $F_n = 1-(\frac{1}{2})^n$, where n is the number of generations of self pollination that have occurred

and a squared inbreeding depression term

$\check{H} = [\sum_i p_i \delta_{ii}]^2$.

These definitions are for a single locus, the equivalent population parameters are simply sums over loci.

The general formula for the covariance between two relatives obtained by self pollination is

$$C_{tgg'} = 2r_{gg'}\sigma_A^2 + \Delta_5\sigma_D^2 + (4\Delta_1+\Delta_2)D_1 + \Delta_1 D_2 + (\Delta_4 - F_g F_{g'} - \Delta_1)\check{H};$$

Table 2. Coefficients of population genetic variances and quadratic components expressed among and within progenies obtained by self fertilization of individuals

Generation	Among					Within				
	σ_A^2	σ_D^2	D_1	D_2	$\check{H}$	σ_A^2	σ_D^2	D_1	D_2	$\check{H}$
S^1	1.0	0.25	1.0	0.125	0	0.5	0.25	1.0	0.375	0.25
S_2	1.5	0.125	2.5	0.563	0.063	0.25	0.125	0.5	0.188	0.125
S_3	1.75	0.063	3.25	0.781	0.047	0.125	0.063	0.25	0.094	0.063
S_4	1.875	0.031	3.63	0.891	0.027	0.063	0.031	0.125	0.047	0.031
S_n	$2F_n$	$\frac{1}{2}(1-F_n)$	$6F_n-2$	$\frac{7}{4}F_n-\frac{3}{4}$	$\frac{1}{2}(2F_n-1)(1-F_n)$	$1-F_n$	$\frac{1}{2}(1-F_n)$	$2(1-F_n)$	$\frac{3}{4}(1-F_n)$	$\frac{1}{2}(1-F_n)$
S_∞	2.0	0	4.0	1.0	0	0	0	0	0	0

assuming

$$\Delta_6=\Delta_7=\Delta_8=\Delta_9=0$$

(Cockerham 1983; see also Jacquard 1974 and Harris 1964).

Genetic variances among and within progenies

The linear function of variances and quadratic components expressed among and within selfed progenies may be derived directly, using the assumed additive, dominance model, with a basic understanding of the effects of self fertilization on the genetic structure of the population. This, however, is tedious, and unnecessary. The genetic variance expressed among a set of progenies can be obtained as the genotypic covariance of individuals within those progenies (Kempthorne 1969). The genetic variance within progenies can be obtained as the difference between the total genetic variance expressed among individuals in a given generation and the genetic variance expressed among the progenies considered.

Breeders of cross pollinated diploid species typically evaluate progenies obtained by self-fertilization of an individual in the previous generation, and employ the S notation. Thus an S_n progeny is obtained by self pollination of a single S_{n-1} individual, and there have been a total of n generations of self pollination. The linear function of population variances and quadratic components expressed among and within S_1 to S_4, S_n, and S_∞ lines are given in Table 2. The corresponding table for the two allele per locus model may be found in Hallauer and Miranda (1981 p. 32).

The coefficients of σ_A^2 both among and within progenies are identical for the two allele and general models. The coefficients of σ_D^2 among and within progenies differs between the two models. In the two allele model the coefficients of σ_D^2 are $F_n/2^n$ and $1-F_n$ among and within progenies, respectively. These coefficients are equal to the sum of the coefficients of σ_D^2 and $\check{H}$ for the general model, as indicated by Cockerham (1983). The coefficients of σ_A^2, D_1 and D_2 among progenies increase linearly with F_n, while the coefficient of $\check{H}$ is quadratic in F_n. The coefficient of σ_D^2 expressed among progenies and all the coefficient for within progenies are linear functions of $1-F_n$. As expected all of the coefficients for within progenies tend toward zero with increased F_n.

The D_1, by definition can be negative. This implies that the variance among S_∞ progenies can be zero in two instances. First, for the trivial case where all variance and quadratic components are zero; and second, the variance among S_∞ progenies can be zero when $D_1=-\frac{1}{2}\sigma_A^2-\frac{1}{4}D_2$. This is the lower limit for D_1.

Breeders of crop species are also interested in the genetic variance expressed among and within selfed progenies which have been bulked for one or more generations. The coefficients of the linear function of population variances and quadratic components expressed among and within these bulked progenies are given in Table 3. The same general trends for the various coefficients which were described for Table 2 are obvious in Table 3 as well.

The extension to include epistatic variances is straightforward. The coefficients on σ_{AA}^2, σ_{AD}^2 and σ_{DD}^2 among S_n progenies would be $(2F_n)^2$, $F_n(1-F_n)$ and $\frac{1}{4}(1-F_n)^2$, respectively and within S_n progenies $1+2F_n-3F_n^2$, $(1-F_n)$ and $3/4(1-F_n)^2$, respectively. The coefficients for the epistatic variances in Table 3 may be obtained as squares or products of the appropriate coefficients.

Predicted selection response

In order to define the necessary regression function, the selection and response units need to be specified. The selection units will be S_n progenies with $n=1,2,3,\ldots,\infty$. The response units will be S_n progenies following recombination of the population and denoted S_n'. The regression function is thus the covariance of S_n and S_n'

Table 3. Coefficients of population genetic variances and quadratic components expressed among and within bulk progenies obtained by self fertilization

Generation in which selfing began	Current generation	Among					Within				
		σ_A^2	σ_D^2	D_1	D_2	$\check{H}$	σ_A^2	σ_D^2	D_1	D_2	$\check{H}$
0	1	1.0	0.25	1.0	0.125	0	0.5	0.25	1.0	0.375	0.25
	2	1.05	0.063	1.5	0.281	0	0.75	0.187	1.5	0.469	0.188
	3	1.0	0.016	1.75	0.383	0	0.875	0.109	1.75	0.492	0.109
	4	1.0	0.004	1.875	0.439	0	0.938	0.059	1.875	0.499	0.059
	5	1.0	0.001	1.938	0.469	0	0.969	0.03	1.938	0.5	0.03
	6	1.0	0	1.969	0.484	0	0.984	0.016	1.969	0.5	0.016
	∞	1.0	0	2.0	0.5	0	1	0	2.0	0.5	0
1	2	1.5	0.125	2.5	0.563	0.063	0.25	0.125	0.5	0.188	0.125
	3	1.5	0.031	2.75	0.641	0.016	0.375	0.094	0.75	0.234	0.093
	4	1.5	0.008	2.875	0.691	0.004	0.438	0.055	0.875	0.247	0.055
	5	1.5	0.002	2.938	0.72	0.001	0.469	0.029	0.937	0.249	0.029
	6	1.5	0	2.969	0.735	0	0.484	0.016	0.969	0.249	0.016
	∞	1.5	0	3.0	0.75	0	0.5	0	1	0.25	0
2	3	1.75	0.063	3.25	0.781	0.047	0.125	0.063	0.25	0.094	0.063
	4	1.75	0.016	3.375	0.82	0.012	0.188	0.047	0.375	0.118	0.047
	5	1.75	0.004	3.438	0.846	0.003	0.219	0.027	0.437	0.123	0.027
	6	1.75	0.001	3.469	0.86	0.001	0.234	0.015	0.469	0.124	0.014
	∞	1.75	0	3.5	0.875	0	0.25	0	0.5	0.125	0
3	4	1.875	0.031	3.63	0.891	0.027	0.063	0.031	0.125	0.047	0.031
	5	1.875	0.008	3.688	0.91	0.007	0.094	0.023	0.187	0.059	0.023
	6	1.875	0.002	3.719	0.923	0.002	0.109	0.014	0.219	0.061	0.013
	∞	1.875	0	3.75	0.938	0	0.125	0	0.25	0.063	0
4	5	1.938	0.016	3.813	0.945	0.015	0.031	0.016	0.063	0.023	0.016
	6	1.938	0.004	3.844	0.955	0.004	0.046	0.012	0.094	0.029	0.011
	∞	1.938	0	3.875	0.969	0	0.063	0	0.125	0.031	0
5	6	1.969	0.008	3.906	0.973	0.008	0.016	0.008	0.031	0.012	0.008
	∞	1.969	0	3.938	0.984	0	0.031	0	0.063	0.016	0
6	∞	1.984	0	3.969	0.992	0	0.016	0	0.031	0.008	0

progeny means divided by the variance among S_n progenies. The S'_n progeny is related to the S_n progeny through an S_0 individual who had as one parent the S_n progeny, with the other gamete coming at random from the population.

Assuming no environmental covariance of relatives, the covariance of S_n and S'_n progeny means is obtained as the expectation of the product of the coded genotypic values minus the product of the expectation of the coded genotypic values. The coded expectation of the genotypic values of both S_n and S'_n is $F_n \sum_i p_i \delta_{ii}$, where F_n is the inbreeding coefficient following n generations of self pollination. Thus, the product of the expectations is equal to

$$F_n^2 \left[\sum_i p_i \delta_{ii}\right]^2 .$$

The expectation of the product of the coded genotypic values is calculated in two parts, and obtained as the sum of the two parts. As indicated previously S_n progenies are obtained by self pollination of S_{n-1} individuals. The population structure in generation $n-1$ is

$$F_{n-1} \sum_i p_i A_i A_i + (1 - F_{n-1}) \sum_i \sum_j p_i p_j A_i A_j .$$

Consider an individual in generation $n-1$ which has genotype $A_i A_i$, the coded genotypic mean value of S_n progeny from this individual is

$$2\alpha_i + \delta_{ii} .$$

The gametic output of the S_n progeny is equal to the gametic output of the S_{n-1} individual, while the gametic output of the population is

$$\sum_k p_k A_k = \sum_i p_i A_i = \sum_j p_j A_j .$$

The distribution of S_0 progeny having as one parent either the individual $A_i A_i$ or its S_n progeny is

$$\sum_k p_k A_i A_k .$$

The S_n' progeny distribution from these S_0 individuals is

$$F_n\left(\tfrac{1}{2}A_iA_i+\tfrac{1}{2}\sum_k p_kA_kA_k\right)+(1-F_n)\sum_k p_kA_iA_k$$

with coded genotypic value of

$$\alpha_i+\frac{F_n}{2}\delta_{ii}+\frac{F_n}{2}\sum_k p_k\delta_{kk}\,.$$

Now, consider an individual in generation $n-1$ with genotype A_iA_j, the coded genotypic mean value of S_n progeny from this individual is

$$\alpha_i+\alpha_j+\tfrac{1}{4}(\delta_{ii}+\delta_{jj}+2\delta_{ij})\,.$$

The gametic output of the S_n progeny is equal to the gametic output of the S_{n-1} individual, i.e.

$$\tfrac{1}{2}(A_i+A_j)\,.$$

The distribution of S_0 progeny having as one parent either the individual or its S_n progeny is

$$\tfrac{1}{2}\left(\sum_k p_kA_iA_k+\sum_k p_kA_jA_k\right).$$

The S_n' progeny distribution from these S_0 individuals is

$$\tfrac{1}{2}\left[\frac{F_n}{2}\left(A_iA_i+A_jA_j+2\sum_k p_kA_kA_k\right)\right.$$
$$\left.+(1-F_n)\left(\sum_k p_kA_iA_k+\sum_k p_kA_jA_k\right)\right]$$

with coded genotypic value of

$$\tfrac{1}{2}\alpha_i+\tfrac{1}{2}\alpha_j+\frac{F_n}{4}\left(\delta_{ii}+\delta_{jj}+2\sum_k p_k\delta_{kk}\right).$$

The covariance of S_n and S_n' progenies is thus

$$F_{n-1}\sum_i p_i(2\alpha_i+\delta_{ii})\left(\alpha_i+\frac{F_n}{2}\delta_{ii}+\frac{F_n}{2}\sum_k p_k\delta_{kk}\right)$$
$$+(1-F_{n-1})\sum_i\sum_j p_ip_j[\alpha_i+\alpha_j+\tfrac{1}{4}(\delta_{ii}+\delta_{jj}+2\delta_{ij})]$$
$$\cdot\left[\tfrac{1}{2}\alpha_i+\tfrac{1}{2}\alpha_j+\frac{F_n}{4}\left(\delta_{ii}+\delta_{jj}+2\sum_k p_k\delta_{kk}\right)\right]$$
$$-F_n^2\left[\sum_i p_i\delta_{ii}\right]^2$$
$$=F_{n-1}\sum_i p_i\left[2\alpha_i^2+(1+F_n)\alpha_i\delta_{ii}\right.$$
$$\left.+F_n\alpha_i\sum_k p_k\delta_{kk}+\frac{F_n}{2}\delta_{ii}^2+\frac{F_n}{2}\delta_{ii}\sum_k p_k\delta_{kk}\right]$$
$$+(1-F_{n-1})\sum_i\sum_j p_ip_j\left[\tfrac{1}{2}\alpha_i^2+\alpha_i\alpha_j+\tfrac{1}{2}\alpha_j^2\right.$$
$$+\frac{2(F_n)+1}{8}\alpha_i\delta_{ii}+\frac{2(F_n)+1}{8}\alpha_j\delta_{jj}$$
$$+\frac{2(F_n)+1}{8}\alpha_i\delta_{jj}+\frac{2(F_n)+1}{8}\alpha_j\delta_{ii}$$
$$+\tfrac{1}{4}\alpha_i\delta_{ij}+\tfrac{1}{4}\alpha_j\delta_{ij}+\frac{F_n}{2}\alpha_i\sum_k p_k\delta_{kk}$$
$$+\frac{F_n}{2}\alpha_j\sum_k p_k\delta_{kk}+\frac{F_n}{16}\delta_{ii}^2+\frac{F_n}{8}\delta_{ii}\delta_{jj}$$
$$+\frac{F_n}{16}\delta_{jj}^2+\frac{F_n}{8}\delta_{ii}\delta_{ij}+\frac{F_n}{8}\delta_{jj}\delta_{ij}+\frac{F_n}{8}\delta_{ii}\sum_k p_k\delta_{kk}$$
$$\left.+\frac{F_n}{8}\delta_{jj}\sum_k p_k\delta_{kk}+\frac{F_n}{4}\delta_{ij}\sum_k p_k\delta_{kk}\right]$$
$$-F_n^2\left[\sum_i p_i\delta_{ii}\right]^2$$
$$=F_{n-1}\left[\sigma_A^2+(1+F_n)D_1+\frac{F_n}{2}D_2+F_n\check{H}\right]+(1-F_{n-1})$$
$$\cdot\left[\tfrac{1}{2}\sigma_A^2+\left(\frac{F_n}{2}+\tfrac{1}{4}\right)D_1+\frac{F_n}{8}D_2+\frac{F_n}{2}\check{H}\right]$$
$$=F_n\sigma_A^2+(F_n^2+\tfrac{3}{2}F_n-\tfrac{1}{2})D_1+\left(\tfrac{3}{4}F_n^2-\frac{F_n}{4}\right)D_2\,.$$

The coefficients for the three components for several levels of inbreeding are given in Table 4.

The regression of S_n' progeny means on S_n progeny means can now simply be defined as

$$\frac{\mathrm{Cov}(S_n,S_n')}{\mathrm{Var}(S_n)}$$

where $\mathrm{Cov}(S_n,S_n')$ has been defined and $\mathrm{Var}(S_n)$ is the variance of S_n progeny means and is equal to

$$\frac{\sigma_e^2}{rE}+\frac{\sigma_{ge}^2}{E}+\sigma_g^2\,,$$

where σ_e^2 is the error variance, σ_{ge}^2 is the genotype × environment interaction variance, and σ_g^2 is the genetic variance expressed among S_n progenies and r and E are the numbers of replications and environments for evaluation, respectively. The error variance can further be expressed as

$$\sigma_e^2=(\sigma_{we}^2+\sigma_{wg}^2)/m+\sigma^2\,,$$

Table 4. Coefficients of population variances or quadratic components involved in the covariance of S_n S_n' progenies

Covariance	σ_A^2	D_1	D_2
S_1, S_1'	0.5000	0.5000	0.0625
S_2, S_2'	0.7500	1.1875	0.2344
S_3, S_3'	0.8750	1.5781	0.3555
S_4, S_4'	0.9375	1.7852	0.4248
S_n, S_n'	F_n	$(F_n^2-\tfrac{1}{2}+\tfrac{3}{2}F_n)$	$\left(\tfrac{3}{4}F_n^2-\frac{F_n}{4}\right)$
S_∞, S_∞'	1.0000	2.0000	0.5000

where σ^2_{we} and σ^2_{wg} are within plot environmental and genetic variances respectively, σ^2 is the plot to plot variance and m is the number of plants per plot. From Table 2, σ^2_{wg} is equal to

$$\sigma^2_{wg} = (1 - F_n)\,[\sigma^2_A + \tfrac{1}{2}\sigma^2_D + 2\,D_1 + \tfrac{3}{4}D_2 + \tfrac{1}{2}\check{H}]\,,$$

and

$$\begin{aligned}\sigma^2_g = {} & 2\,F_n\,\sigma^2_A + \tfrac{1}{2}(1 - F_n)\,\sigma^2_D + (6\,F_n - 2)\,D_1 \\ & + (\tfrac{7}{4}F_n - \tfrac{3}{4})\,D_2 + \tfrac{1}{2}(2\,F_n - 1)\,(1 - F_n)\,\check{H}\,.\end{aligned}$$

When recombination takes place only among selected progenies, then, the regression function is the regression of S_n progeny mean on mid-parent value, the numerator is unchanged, but the variance of the mid-parent value is $\frac{1}{2}$ the variance of the S_n progeny mean. Thus, predicted gains can be expressed as

$$\begin{aligned} G = (\bar{X}_{ssn} - \bar{X}_{sn}) & \Bigg[F_n\,\sigma^2_A + (\tfrac{3}{2}F_n + F^2_n - \tfrac{1}{2})\,D_1 \\ & + \left(\tfrac{3}{4}F^2_n - \frac{F_n}{4}\right) D_2 \Bigg] \Big/ \tfrac{1}{2} - F^2_n\,\check{H} \\ \cdot \Bigg[& \sigma^2_{we} + (1 - F_n)\,[\sigma^2_A + \tfrac{1}{2}\sigma^2_D + 2\,D_1 + \tfrac{3}{4}D_2 + \tfrac{1}{2}\check{H}]/_{mre} \\ & + \frac{\sigma^2}{rE} + \frac{\sigma^2_{ge}}{E} + 2\,F_n\,\sigma^2_A + \tfrac{1}{2}(1 - F_n)\,\sigma^2_D \\ & + (6\,F_n - 2)\,D_1 + (\tfrac{7}{4}F_n - \tfrac{3}{4})\,D_2 \\ & + \tfrac{1}{2}(2\,F_n - 1)\,(1 - F_n)\,\check{H} \Bigg]\,, \end{aligned}$$

where all terms are defined as previously and $\bar{X}_{ssn}$ and $\bar{X}_{sn}$ are the means of the selected S_n progeny and the population of S_n progenies, respectively.

The formulae for G, where the assumption of two alleles per locus is made, have in the numerator some fraction of the additive genetic variance, i.e., gains are directly proportional to the additive variance in the population. In contrast, when this assumption is relaxed, selection response involves not only the additive genetic variance, but also D_1 and D_2. The D_1 can be negative as noted by Cockerham (1983). This suggests that even for populations with larger values for σ^2_A, predicted gains may be low. In particular $\mathrm{Cov}(S_\infty, S'_\infty)$, i.e., doubled haploids, is equal to $\sigma^2_A + 2\,D_1 + \frac{1}{2}D_2$ which, if D_1 takes its lowest possible value, will be equal to zero. Thus, despite a large value for σ^2_A, no selection response may be realized using doubled haploids under these circumstances. Additionally, if D_1 is negative, selection response will, in most instances, be less than predicted with the assumption of two alleles per locus. Use of models which assume two alleles per locus to predict gains for populations with multiple alleles is inappropriate, and may give inflated estimates of selection response.

Some suggested progenies to evaluate in estimation of population variances and quadratic components

With 5 population variances or quadratic components to estimate we need a minimum of 5 genetic variances or covariances which are linear functions of these population parameters. The mechanism for obtaining the proposed progenies is as follows: a random set of S_0 individuals is self pollinated to obtain a set of S_1 lines; equal size random sets of S_1 lines are designated males and females; a Design II (Comstock and Robinson 1948; Hallauer and Miranda 1981) mating is produced by crossing S_1 lines and a set of individuals within each S_1 line is also self pollinated. The progenies produced for evaluation are the set of full sib families from the Design II, the S_1 lines and the S_2 bulk progenies.

From the analysis of the Design II both σ^2_A and σ^2_D can be estimated. From the analysis of the S_1 lines the component for the genetic variance among S_1 lines is equal to

$$\sigma^2_A + \tfrac{1}{4}\sigma^2_D + D_1 + \tfrac{1}{8}D_2\,.$$

From the analysis of the S_2 bulks, the component for the genetic variance among S_2 bulks is equal to

$$\tfrac{3}{2}\sigma^2_A + \tfrac{1}{16}\sigma^2_D + \tfrac{3}{2}D_1 + \tfrac{9}{32}D_2\,.$$

Thus far 5 linear functions in 4 unknowns have been defined, hence each can be uniquely estimated and their standard errors defined (see Hallauer and Miranda 1981).

At least three genotypic covariances among these progenies can be defined; the covariance of an inbred progeny mean with half sib family mean where the S_0 individual is common can be shown to be

$$\mathrm{Cov}\,(S_n, \mathrm{H.S.}) = \tfrac{1}{2}\sigma^2_A + D_1\,.$$

The covariance of full sib progeny means with the average of the self progeny means of the two parents is

$$\mathrm{Cov}\,(\mathrm{F.S.}, \bar{S}_1) = \tfrac{1}{2}\sigma^2_A + \tfrac{1}{4}D_1\,.$$

This second covariance is identical to $\mathrm{Cov}\,(\mathrm{H.S.}, S_1)$. Finally, the covariance of S_1 line mean with S_2 bulk progeny mean is;

$$\mathrm{Cov}\,(S_1, S_2) = \sigma^2_A + \tfrac{1}{8}\sigma^2_D + \tfrac{5}{4}D_1 + \tfrac{3}{16}D_2 - \tfrac{3}{8}\check{H}\,.$$

Thus there exists a set of up to nine linear functions of the 5 population parameters, and the corresponding point estimates can be obtained by least squares. These progenies in contrast to those proposed by Cockerham (1983) may provide more information on the dominance variance.

This completes the extension of theory from 2 alleles per locus to multiple alleles per locus as it applies to both variances among and within progenies and selection response. Hopefully this work clarifies for some the condensed coefficients identity which at times are complex. This work complements the results of Matzinger and Cockerham (1985) who didn't address this particular type of selection response, but dealt with some closely related and equally important areas.

References

Chevalet C, Gillois M (1977) In: Pollak E, Kempthorne O, Bailey TB, Jr (eds): Proc Int Conf Quant Genet. Iowa State University Press, Ames, Iowa

Cockerham CC (1969) Notes on quantitative genetics. North Carolina State University, Raleigh NC

Cockerham CC (1983) Covariances of relatives from self-fertilization. Crop Sci 23:1177–1180

Cockerham CC, Matzinger DF (1985) Selection response based on selfed progenies. Crop Sci 25:483–488

Cowen NM, Weber GP, Frey KJ (1985) BASIC programs for estimating coefficients of inbreeding, parentage, and identity. J Hered 76:401

Gillois M (1964) La relation d'identité génétique. Thesis, Faculté des Sciences, Paris

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

Harris DL (1964) Genotypic covariance between inbred relatives. Genetics 50:1319–1348

Horner TW, Weber CR (1956) Theoretical and experimental study of self-fertilized populations. Biometrics 12:404–414

Jacquard A (1974) The genetic structure of populations (transl by Charlesworth D and Charlesworth B). Springer, Berlin Heidelberg New York

Kempthorne O (1969) An introduction to genetic statistics. Iowa State University Press, Ames, Iowa

Nyquist WE (1978) Notes on statistical genetics. Purdue University, West Lafayette Ind

Sprague GF (1966) Quantitative genetics in plant improvement. In: (ed) Frey KJ: Plant breeding. Iowa State University Press, Ames, Iowa, pp 5–10

Weir BS, Cockerham CC (1977) Two-locus theory in quantitative genetics. In: Pollak E, Kempthorne O, Bailey TB Jr (eds): Proc Int Conf Quant Genet. Iowa State University Press, Ames, Iowa, pp 247–269