Theoretical and Applied Genetics 45, 163–169

A Genetic Theory of General Varietal Ability for Diploid Crops

A. J. WRIGHT

Plant Breeding Institute, Cambridge (England)

Summary. The use of estimates of general varietal ability (g.v.a.) of individual plants, as defined in an earlier paper (Wright, 1973), is examined in the contexts of both the population improvement and synthetic variety building phases of a breeding programme, and its theoretical efficiency compared with those of parental and g.c.a. methods of assessment. It is concluded that selection based on g.v.a. may be useful during the varietal building phase when heritability is low. In the absence of epistasis, the g.v.a. variance is shown to account for all the variance among the synthetics of size s which can be drawn from a large random-bred population, except for $\frac{(s-1)}{2s^3}\sigma_D^2$. The possibilities of the prediction of synthetic variety performance from I_1 and polycross progeny data is discussed.

1. Introduction

In a recent publication (Wright, 1973) it was argued that although estimates of general combining ability (g.c.a.) were appropriate for the improvement of a large population or the first generation of synthetic varieties, the subsequent change due to inbreeding during multiplication should be taken into account by any method used to assess potential varietal constituents. The general varietal ability (g.v.a.) of an individual for diploid synthetics of a given size (s) was defined as the mean of all synthetics of this size having the individual as a common parent which could be drawn from a population of n. By expressing the expectation of any synthetic as the mean of all possible F_1 and I_1 progenies among its parents, it was shown that the g.v.a. of any individual can be expressed as a combination of its g.c.a. and the mean of its I_1 (first generation inbred) progeny, the latter weighed according to the size of population and synthetic.

The present paper examines the genetic basis of general varietal ability and its potential utility during the population improvement and varietal building phases of a breeding programme and compares it with more conventional means of assessment in these contexts. The possibilities of prediction of varietal performance from parental g.v.as. is also discussed.

2. The Definition of General Varietal Ability

The genetic model used throughout this paper is based on a single gene under disomic inheritance with beneficial allele A at population frequency p and allelomorph a at frequency $q(=1-p)$. The letters u and v are used to symbolise the frequencies of alleles A and a in synthetic varieties, and r, z, and w are as defined in the text. Genotypes AA, Aa and aa are given values a, d, and $-a$ respectively, where d takes positive or negative sign according to the direction of dominance. The absence of epistasis and linkage, and panmixis without gene immigration or natural selection in varietal multiplication plots, are assumed.

The derivation of a progeny testing method for the direct estimation of general varietal ability was given in a previous paper (Wright, 1973), and will be summarised here. Defining the general varietal ability of an individual (the ith) as the mean of all synthetics of size s which have this individual as one of s parents chosen from a population of n, its expection can be expressed in terms of the $n^2 I_1$ and F_1 families among these n parents as

$$\frac{1}{s^2}a_{ii} + \frac{2(s-1)}{2s^2(n-1)}\Sigma a_{ik} + \frac{(s-1)}{s^2(n-1)}\Sigma a_{kk} + \\ + \frac{(s-1)(s-2)}{s^2(n-1)(n-2)}\Sigma a_{kl}$$

where a_{ii} refers to the inbred progeny of the ith parent, a_{ik} to F_1 progenies involving this parent, a_{kk} to all other inbred progenies, and a_{kl} to all other F_1 families. Subtracting from this the sum of all I_1 families multiplied by a factor of $\frac{(s-1)}{s^2(n-1)}$ and of all F_1 families multiplied by $\frac{(s-1)(s-2)}{s^2(n-1)(n-2)}$ (i.e. terms which are common to all parents and therefore without any effect on comparisons among g.v.as.) leaves

$$\frac{(n-s)}{s^2(n-1)}a_{ii} + \frac{(n-s)(s-1)}{s^2(n-1)(n-2)}\Sigma a_{ik}.$$

Since Σa_{ik} is equal to $2(n-1)$ times the mean of F_1 progenies of the ith parent, then

$$\text{g.v.a.}_i = \frac{(n-s)}{s^2(n-1)}\bar{I}_{1i} + \frac{2(n-s)(s-1)}{s^2(n-2)}\text{g.c.a.}_i, \qquad (1)$$

$$\text{or} \quad \frac{1}{s^2}\bar{I}_{1i} + \frac{2(s-1)}{s^2}\text{g.c.a.}_i \text{ when } n \text{ is large relative to } s \qquad (2)$$

where g.c.a. can be estimated as the mean of F_1 progenies or the polycross progeny mean.

This definition of g.v.a. in terms of g.c.a. and I_1 progeny mean not only suggests a means of estimation in practice, but also facilitaties its quantification in terms of the genetic symbols already defined.

3. Population Improvement with a View to Synthetic Variety Production

In the improvement of a large population, the aim of the progeny test, or other means of assessment, is to evaluate the genes carried by an individual with respect to their potential effect on the mean of the selected offspring generation. Any gain in performance of this new population over the unselected one will depend directly on improvements of the gene frequencies at some or all loci concerned with the expression of the character. The relationship between population mean and gene frequency is not a linear one, but in general the size of the population and of the selected fraction is sufficiently large for the effect on gene frequency of the inclusion or rejection of any individual to be very small. Hence, for the small range of gene frequency concerned, no serious error is involved by the assumption that its relationship with population mean is linear. This assumption is basic to the normal theories of selection and response (see Falconer, 1960). The choice of genotype AA in preference to Aa is therefore equally beneficial as the choice of Aa in preference to aa, since the effect on the population gene frequency will be the same. Hence, only that portion of the genotypic variance concerned with the linear effects of a gene substitution is utilisable by selection (e.g. Fisher, 1930).

Although selection is normally carried out with the aim of improving the mean of the population itself, the above discussion applies equally to selection aimed at the improvement of the potential of the population to produce good parents for synthetic varieties, a quantity which can be more precisely defined as the mean of all synthetics of a certain size which could be drawn from the population.

In the construction of formulae for the prediction of response to selection based on any method of assessment, the process can usefully be broken down into two components, allowing a fuller understanding of the factors affecting the efficiency of selection. These are the change in gene frequency brought about by selection, and the consequent change in population performance. The effect of one generation of selection of the population frequency (p) of allele A can be expressed in terms of the linear regression of p onto the assessment values (x) on which selection of intensity i is based as

$$\Delta_p = i\sigma_x b_{p\cdot x} = i\sigma_{p\cdot x}/\sigma_x \qquad (3)$$

where $\sigma_{p\cdot x}$ is the covariance of p and x (Falconer, 1960). Turning to the second component of response, it is first necessary to quantify the mean of synthetics which could be drawn from the population. The fragmentation of the population into a number of smaller random mating groups (i.e. synthetic varieties) incurs inbreeding with a coefficient of $1/2s$ (Falconer, 1960). The mean of all synthetics of size s which can be produced is therefore

$$y_s = (p - q)\, a + 2pq\left(1 - \frac{1+F}{2s}\right) d, \text{ (see Gallais, 1967).}$$

The rate of change of this expression with respect to a change in gene frequency is found most simply by differentiation as

$$\frac{\mathrm{d}y_s}{\mathrm{d}p} = 2\left\{a + d(q - p)\left(1 - \frac{1+F}{2s}\right)\right\} = 2\alpha_{FS} \qquad (4)$$

The expected response to selection, in terms of synthetic variety production, is therefore the product of expressions (3) and (4), after summation over genes, with the appropriate expectations for $\sigma_{p\cdot x}$ and σ_x substituted. Table 1 gives the expectation of these statistics when selection is based on parental phenotype or estimates of general combining ability or general varietal ability, and finally the expected response to selection of unit intensity. It should be noted that the error variance (σ^2_E) has been included only for the sake of completeness, and cannot be regarded as necessarily being constant for all methods of assessment used, variation being introduced according to the field design used (especially in the case of g.v.a. estimation), as well as by possible differential within family variances and genotype × environment interaction. For this reason, factors of s are of little significance here, and may be partially cancelled from the formulae. The right hand column of the table gives expectations for the case $F = 0$, where $\sigma^2_{AS} = \sum 2uv\alpha^2_s$ and $\sigma_{A\cdot AS} = \sum 2uv\alpha\alpha_s$.

The occurrence of the term α_s in both (1) and (2) for selection in a panmictic population based on g.v.a., and resulting in the appearance of α^2_s terms in the response formula demonstrates an important property of this means of assessment. As total response is derived by summation over all genes, the selective effort expended in the advancement of any gene is proportional to the rate of response of that gene in terms of synthetic variety improvement. Clearly, the selective effort generated by the differentiation among assessment values is applied as efficiently as possible. A strict analogy exists between this situation and that prevailing during selection for population mean *per se* based on breeding value or g.c.a. with α_s replaced by α. It is clear from the table that the efficiency of g.c.a. or parental assessment will be improved as the correlation of α and α_s over all genes is increased. This correlation will be increased not only in the trivial cases where dominance levels are low or s is high, but also when α and hence α_s values are high for all or most genes; that is when additive variance is high and narrow sense heritability consequently large. Apart from these observations, it is not possible to say what are the relative efficien-

Table 1. *Improvement in population potential for synthetic variety production following selection of unit intensity based on various criteria*

	General case	$F = 0$
Variances of:		
phenotype	$\sum 2pq\,(1+F)\,\alpha_F^2 + \sum 4p^2q^2\,(1-F^2)\,d^2 + \sigma_E^2$	$\sigma_A^2 + \sigma_D^2 + \sigma_E^2$
g.c.a.	$\sum \frac{1}{2} pq\,(1+F)\,\alpha^2 + \sigma_E^2$	$\frac{1}{4}\sigma_A^2 + \sigma_E^2$
g.v.a.	$\frac{\sum 2pq\,(1+F)\,\alpha_{FF}^2}{s^2} + \frac{\sum p^2q^2\,(1-F^2)\,d^2 + \sigma_E^2}{s^4}$	$\frac{1}{s^2}\sigma_{AS}^2 + \frac{1}{4s^4}\sigma_D^2 + \sigma_E^2$
Covariances of p with:		
phenotype	$\sum pq\,(1+F)\,\alpha_F$	$\sum pq\alpha$
g.c.a.	$\sum \frac{1}{2} pq\,(1+F)\,\alpha$	$\sum \frac{1}{2} pq\alpha$
g.v.a.	$\frac{\sum pq\,(1+F)\,\alpha_{FF}}{s}$	$\frac{\sum pq\alpha_S}{s}$
Response to selection for:		
phenotype	$\frac{\sum 2pq\,(1+F)\,\alpha_F\alpha_{FS}}{\sqrt{\sum 2pq\,(1+F)\,\alpha_F^2 + \sum 4p^2q^2\,(1-F^2)\,d^2 + \sigma_E^2}}$	$\frac{\sigma_{A\cdot AS}}{\sqrt{\sigma_A^2 + \sigma_D^2 + \sigma_E^2}}$
g.c.a.	$\frac{\sum pq\,(1+F)\,\alpha\alpha_{FS}}{\sqrt{\sum \frac{1}{2} pq\,(1+F)\,\alpha^2 + \sigma_E^2}}$	$\frac{\frac{1}{2}\sigma_{A\cdot AS}}{\sqrt{\frac{1}{4}\sigma_A^2 + \sigma_E^2}}$
g.v.a.	$\frac{\sum 2pq\,(1+F)\,\alpha_{FF}\alpha_{FS/s}}{\sqrt{\sum 2pq\,(1+F)\,\alpha_{FF/s^2}^2 + \sum p^2q^2\,(1-F^2)\,d^2/s^4 + \sigma_E^2}}$	$\frac{(1/s)\,\sigma_{AS}^2}{\sqrt{(1/s^2)\,\sigma_{AS}^2 + (1/4s^4)\,\sigma_D^2 + \sigma_E^2}}$

where: $\alpha = a + d(q-p)$

$\alpha_S = a + d\,(q-p)\,(1 - 1/2s)$

$\alpha_F = a + d\,(q-p)\,\frac{(1-F)}{(1+F)}$

$\alpha_{FF} = a + d\,(q-p)\left(\frac{(s-1)}{s} + \frac{(1-F)}{(1+F)\,2s}\right)$

$\alpha_{FS} = a + d\,(q-p)\left(1 - \frac{1+F}{2s}\right)$

Table 2. *The relative efficiencies of varietal building methods*

	General case	$F = 0$
Covariance of g.v.a. with:		
phenotype	$\frac{\sum 2pq\,(1+F)\,\alpha_F\alpha_{FF}}{s} + \frac{\sum 2p^2q^2\,(1-F^2)\,d^2}{s^2}$	$\frac{1}{s}\sigma_{A\cdot AS} + \frac{1}{2s^2}\sigma_D^2$
g.c.a.	$\frac{\sum pq\,(1+F)\,\alpha\alpha_{FF}}{s}$	$\frac{1}{2s}\sigma_{A\cdot AS}$
Response to selection for:		
phenotype	$\frac{\sum 2pq\,(1+F)\,\alpha_F\alpha_{FF} + \sum 2p^2q^2\,(1-F^2)\,d^2/s}{\sqrt{\sum 2pq\,(1+F)\,\alpha_F^2 + \sum 4p^2q^2\,(1-F^2)\,d^2 + \sigma_E^2}}$	$\frac{\sigma_{A\cdot AS} + (1/2s)\,\sigma_D^2}{\sqrt{\sigma_A^2 + \sigma_D^2 + \sigma_E^2}}$
g.c.a.	$\frac{\sum pq\,(1+F)\,\alpha\alpha_{FF}}{\sqrt{\sum \frac{1}{2} pq\,(1+F)\,\alpha^2 + \sigma_E^2}}$	$\frac{\frac{1}{2}\sigma_{A\cdot AS}}{\sqrt{\frac{1}{4}\sigma_A^2 + \sigma_E^2}}$
g.v.a.	$\frac{\sum 2pq\,(1+F)\,\alpha_{FF/s}^2 + \sum p^2q^2\,(1-F^2)\,d^2/s^3}{\sqrt{\sum 2pq\,(1+F)\,\alpha_{FF/s^2}^2 + \sum p^2q^2\,(1-F^2)\,d^2/s^4 + \sigma_E^2}}$	$\frac{(1/s)\,\sigma_{AS}^2 + (1/4s^3)\,\sigma_D^2}{\sqrt{\frac{1}{s^2}\sigma_{AS}^2 + \frac{1}{4s^4}\sigma_D^2 + \sigma_E^2}}$
Among synthetics	—	$\frac{(1/s)\,\sigma_{AS}^2 + (2n - 1/4s^3)\,\sigma_D^2}{\sqrt{\frac{1}{s}\sigma_{AS}^2 + \frac{2n-1}{4s^3}\sigma_D^2 + \sigma_E^2}}$

cies of the various means of assessment, if only because of the inconsistency of σ^2_E, but, the error of estimation of g.v.a. will almost inevitably be large, lowering the efficacy of this method.

However, some expected results of hypothetical selection for g.v.a., particularly in the longer term, may be examined. One result of the optimal balance between potential response and applied selection pressure already discussed is that no selection is applied to any gene which is at its optimal population frequency. This, is identical with the case of normal selection for population mean, except that an equilibrium gene frequency can now only exist when

$$d\left(1-\frac{1}{2s}\right) > a$$

requiring a higher level of overdominance as the size of synthetic is reduced. Selection for g.v.a. will favour the recessive homozygote when gene frequency exceeds this equilibrium, that is

$$p > \frac{a + d\left(1-\frac{1}{2s}\right)}{2d\left(1-\frac{1}{2s}\right)}.$$

It can also be seen that for any value of s, when d is positive α_s is greater than α when p is greater than 0.5, but smaller when p is less. Hence selection for g.v.a. will apply a relatively greater selection pressure on genes for which $p > 0.5$, and less when $p < 0.5$ than would any conventional method of assessment. The reverse will be true for genes for which d is negative. Over a series of selection cycles, this would therefore tend to bring about a different disposition of gene frequencies, with genes with positive dominance increments tending more towards extremely low or high frequencies, and those with negative dominance more towards intermediate frequencies. This is a reflection of the relatively lower importance of maintaining heterozygosity which cannot be subsequently utilised by synthetic varieties, and in its differential effects on genes with positive and negative dominance relations shows interesting similarities to the theory of reciprocal recurrent selection.

4. Building the Synthetic

Although, as already shown, the relationship between gene frequency and performance can be assumed linear when selection is aimed at population improvement, this is not the situation during the varietal building phase of the breeding programme. The effect on varietal gene frequency of the selection or rejection of any individual is larger, particularly with small synthetics, and can involve such a large change in gene frequency that the curvature of the relationship, which is the basis of inbreeding depression, must be taken into account.

In considering the role of the use of g.v.a. in this context, comparisons among the three genotypes AA, Aa, and aa will be made in respect both of their g.v.a. values, and the means of all synthetics they enter into. The relevant contrasts to be made are linear $((AA-aa)/2)$ and quadratic $(Aa-(AA+aa)/2)$, and hence the ratio of these two, giving a measure of the curvature of the relationship, in what might be regarded as a form of 'dominance' ratio. It must be noted that in the derivation of comparisons among synthetic means, the value of a synthetic with mean gene frequency will be used in place of the mean of synthetics with varying gene frequencies. Although these quantities are not the same, the purely quadratic nature of the relationship of gene frequency and variety mean ensures that the difference is constant for a given value of s, irrespective of the mean gene frequencies involved, and so the discrepancy disappears on making the comparisons.

(a) *For large n*

When $(s-1)$ parents for a synthetic have been chosen from the population, the mean of the resultant synthetic will be

$$y_s = (u-v)a + 2uvd\,,$$

where

$$u = \frac{(s-1)\,p + z}{s}$$

where p is the gene frequency of the population and hence the mean frequency of the $(s-1)$ chosen parents, and z that of the sth. Giving z values of 1, $\frac{1}{2}$ or 0, the following relations can easily be found among the three possible synthetics:

$$\text{linear} = \frac{a + d(q-p)\left(1-\frac{1}{s}\right)}{s}$$

$$\text{quadratic} = \frac{\frac{1}{2}d}{s^2}$$

$$\text{and thus quadratic/linear} = \frac{d}{2s\left(a + d(q-p)\left(1-\frac{1}{s}\right)\right)}. \tag{5}$$

If similar comparisons are made among g.v.a. values, linear and quadratic contrasts are

$$s\left(a + d(q-p)\left(1-\frac{1}{2s}\right)\right)\Big/s^2$$

and

$$\frac{1}{2d}\Big/s^2$$

leading to a ratio identical with that at (5).

(b) *For small n*

For the case where n is small relative to s, the situation is changed in three respects. First, the exact contribution of the I_1 progeny mean to the modified progeny test, as in formula (1), must be considered. Secondly each genotype represented in the polycross

receives a unique pollination, since the $(n-1)$ genotypes contributing to it are significantly different from any other set of $(n-1)$. Thirdly, the $(n-1)$ genotypes from which $(s-1)$ may be chosen as companions to the sth also vary from genotype to genotype.

In making comparisons between any two individuals with respect to the mean of all synthetics they can enter, some synthetics involve both individuals and contribute nothing to the comparison. Of the remainder, each synthetic containing the first can be matched with one containing the second but otherwise identical. The gene frequency of the remaining $(n-2)$ genotypes all of which contribute equally to this second class of synthetics is therefore

$$r = \frac{np - w - z}{(n-2)}$$

where w and z are the gene frequencies of the two genotypes being compared. A contrast between any two synthetics is given by

$$y_1 - y_2 = \frac{2}{s}\frac{\{a(w-z) + d\{(w-z)\{s-2(s-1)r\} - w^2 + z^2\}\}}{s}.$$

Substituting values of 1, $\frac{1}{2}$ and 0 as appropriate, the linear and quadratic contrasts among genotypes are found to be

$$\frac{\left\{(n-2)\,a + d(q-p)\,n\left(1-\frac{1}{s}\right)\right\}}{s\{n-2\}}$$

and

$$\frac{(n-2s)d}{2s^2(n-2)}$$

giving a "dominance" ratio

$$\frac{(n-2s)d}{2s\left\{(n-2)\,a + d(q-p)\,n\left(1-\frac{1}{s}\right)\right\}}. \qquad (6)$$

Turning to the expectations of g.v.a. (equation 1) in this situation, as noted earlier, each individual receives pollen from the remaining $(n-1)$ in the population. Whether estimated as a polycross mean or by controlled pair crossing, its g.c.a. value is therefore equivalent to the mean of only $(n-1)$ F_1 progenies. In terms of the single locus model discussed here, the usual biometrical expectation of g.c.a. value when n is large can be restored by means of the inclusion of the inbred progeny mean with a coefficient of $1/n$. Hence,

$$\frac{(n-1)\,\bar{F}_1 + \bar{I}_1}{n} = \text{g.c.a.}$$

$$\therefore\ \bar{F}_1 + \bar{I}_1/(n-1) = \text{g.c.a.}\left(\frac{n}{(n-1)}\right)$$

and the g.v.a. value is therefore

$$\frac{2(n-1)(s-1)\,\bar{F}_1}{s^2(n-1)} + \frac{(n-2)}{s^2(n-1)}\,I_1 = \text{g.c.a.}\,\frac{2(s-1)}{s^2(n-1)} + \frac{(n-2s)}{s^2(n-1)}\,I_1.$$

The coefficient in the last term is therefore the residual weighting applied to the I_1 progeny mean after restoration of the biometrical expectation of g.c.a. for a large population. Comparing the values of the modified progenies for the three genotypes then gives linear and quadratic contrasts:

$$\frac{s\left\{(n-2)\,a + d(q-p)\,n\left(1-\frac{1}{s}\right)\right\}}{s^2(n-1)}$$

and

$$\frac{(n-2s)\,\frac{1}{2d}}{s^2(n-1)}$$

and a "dominance" ratio identical with (6).

It is clear, therefore, that irrespective of the size of the population under test, general varietal ability applies a bias in favour of heterozygous genotypes which matches the advantage of synthetics with these included as parents, and is therefore optimal in terms of ranking individuals as potential parents.

5. Relative Efficiencies of Varietal Building Methods

The variance among g.v.a. values as given in Table 1 represents the variance among groups of synthetics with one common parent, and hence the variation accounted for by fitting one g.v.a. to each synthetic. Therefore, fitting the g.v.a. values of all s independent parents accounts (when $F = 0$) for

$$\Sigma\frac{2pq\,\alpha_S^2}{s} + \frac{1}{4s^3}\sigma_D^2. \qquad (7)$$

The demonstration in the foregoing Section that all variation among g.v.a. values is utilisable during variety building confirms that all the variance of g.v.a. values is part of the total variance among synthetics. This allows the expected response to selection of varietal parents to be expressed simply as the ratio of the variance of g.v.a. values to the standard deviation of their estimated values. This expectation, after some cancellation of terms in s, is given in Table 2. By the application of conventional theory of correlated responses (Falconer, 1960), the expected response to selection of varietal parents using estimates of g.c.a. or phenotype can now be found as

$$i\frac{\sigma_{x\cdot g}}{\sigma_x}(s)$$

where $\sigma_{x\cdot g}$ is the covariance of the assessment criterion with g.v.a. The expectation of these covariances are given in the upper part of Table 2 and the response formulae below. As in Table 1, no equivalence exists between the different error variances, and definite assertions as to the superiority of any particular method cannot be made. However, the potential advantage of g.v.a. assessment is greater for this varietal building phase than during population improvement, because both linear and non-linear com-

ponents of variance are of direct benefit. Again, the efficiency of conventional methods of assessment will be enhanced by an increased correlation of α and α_s.

Hill (1971) gives the variance among all synthetics which can be drawn from a random bred population as

$$\sum \frac{2pq\,\alpha_s^2}{s} + \frac{(2s-1)}{4s^3}\sigma_D^2 . \qquad (8)$$

Hence, the expected response to selection among synthetics is as shown in Table 2. If the error variance of g.v.a. and synthetic estimation were the same, then the former method would be expected to be slightly the more efficient. However, Hill suggests that selection among all but the smallest synthetics will always be a less efficient procedure than selection of parents, because the variance among such synthetics is low. But it is clear that the success of the varietal building process must always be limited by the quantity of variation available among possible synthetics. As in the case of a hybrid breeding programme, the advantage of parental selection lies in the far smaller number of trial entries which have to be tested, and in addition, the fewer seasons required to produce seed for progeny testing than for second generation synthetics.

In a random bred population, the variance among synthetics not accounted for parental g.v.as. is only $\frac{(s-1)}{2s^3}\sigma_D^2$. This suggests that, at least when the necessary assumptions are fulfilled, the performance of any specific synthetic can be predicted from the g.v.as. of its parents with a high degree of accuracy.

6. Discussion

The examination of the role of general varietal ability in both population improvement and varietal construction phases of a breeding programme demonstrates the essential continuity of the underlying theory. Although the model used in the development has been highly restricted in terms of assumptions with regard to epistasis and multiple allelism, it is probable that these could be considerably relaxed. The chief requirement with regard to epistatic interactions is that the relationship between coefficient of inbreeding and population mean should be linear, or very nearly so, for all possible synthetics which can be produced. Multiple alleles can probably be dealt with in much the same way as in the treatment of more familiar population parameters, although this might itself depend on the absence of epistasis and linkage (Hill, 1971).

The important difference between the population improvement phase of the breeding programme, whether this is a single cycle of selection and recombination or part of a recurrent series, and the varietal building phase, is that the former is concerned to raise the performance of synthetics which might be produced from the selected group after reconstitution as a random mating population, whereas the latter is concerned simply with the performance of the synthetics to be produced from the selected group as it stands. This is in direct parallel with the more usual types of selection and response prediction where an improvement is required in the mean of the selected population after a generation of random mating, or of the mean of the selected group *per se*. This analogy is strikingly revealed by a comparison of the response formulae in Tables 1 and 2, where those dealing with the improvement of the potential of the following generation (Table 1) resemble response formulae based on narrow-sense heritability, and those concerned with immediate improvement resemble formulae using heritability in the broad sense. In the cases both of population mean and varietal mean improvement, the portion of the immediate gain obtained which is due to the non-linear portion of variance is lost during the subsequent random mating cycle. It is also interesting to note that the analogy is not between g.v.a. and g.c.a. as might have been expected, but between g.v.a. and genotype.

The term α_s can be described as the "average effect" (Fisher, 1930) of a gene in terms of the production of synthetic varieties, whereas the "average excess" is represented by α under conventional assessment and α_s in the case of g.v.a. The equivalence of average excess and effect achieved by the use of g.v.a. is also the rationale behind other breeding procedures, occurring in its most advanced form in reciprocal recurrent selection (Comstock, Robinson, and Harvey, 1949), and leads to a linear component of variance which is always directly utilisable by selection. In the case of the improvement of a large population this is the variance of general combining abilities, which, for hybrid improvement, is defined with respect to the reciprocal population and is a component of variation among crosses in the hybrid population (Griffing, 1962). In the case of synthetic variety production, it has been similarly shown that the linear variance is a component of the total variance among synthetics.

It was suggested earlier that the covariance of α and α_s terms across genes, upon which the efficiency of conventional assessment in part depends, will be high when narrow-sense heritability is high. This, of course, is the very condition under which progeny testing of any sort tends to be less efficient than simpler phenotypic assessment. Therefore, whereas parental phenotypic assessment will tend to be the most efficient procedure early in a breeding programme, the reduction in heritability later on may impose the need for some form of progeny testing. At this stage, the use of g.c.a. estimates will probably still be more effective than g.v.a., if only because of the inherent practical difficulties of the estimation of g.v.a. values for the members of a large breeding population. But it is during the final phase of varietal construction

that the estimate of g.v.a. has its greatest potential. At this time, heritability will probably be low, and because of the relatively low correlation of α and α_s when genes are close to their optimal frequencies, and because of the non-linear component of variance among g.v.a. values, considerable alterations of the ranking of potential parents according to the two methods may occur. Another reason for the preference of g.v.a. at this time is that, providing earlier selection has been successful, most genes will have the beneficial allele at a frequency in the range 0.5 to 1. This means that for those genes with positive dominance relations α_s values will be higher than α, while negative dominance will lead to higher α terms. Since the existence of inbreeding depression itself depends on the predominance of positive dominance, then the linear component of variance of g.v.a. will be greater than the linear 'covariance' of g.c.a., thus enhancing the efficiency of g.v.a. assessment. Of course, the significance of all these arguments will be strengthened as the size of synthetic to be produced is reduced.

In Section (5) it was concluded that, in a large random-bred population, without multiple alleles or epistatic interaction, the variance among synthetics not accounted for by general varietal ability is very small. The important new assumption here concerns the size of the population, since the variance among synthetics from a limited population is unknown, and the significance of the changed coefficients of g.c.a. and I_1 in the expectation of g.v.a. under these conditions is uncertain. However, the implication is that specific varietal effects are likely to be unimportant, and since g.v.a. is a function only of g.c.a. and I_1 progeny, the specific combining abilities of individual crosses among the parents have a trivial effect on the expectation of any varietal mean. Estimation of the importance of specific varietal effects experimentally could be achieved by means of designs such as those suggested by Hill (1966). Although a considerable area of uncertainty remains, which can probably only be resolved by empirical means, the estimation of general varietal ability offers a promising basis for the prediction of the mean performance of any specified synthetic by means of I_1 and polycross progenies or any other direct estimator of g.c.a. values.

A common problem in varietal construction is to decide on the optimal number of parents. Each additional parent has a lower genetic potential than those already included, but decreases the loss due to inbreeding (Graumann, 1952). In quantitative terms, this is expressed as an increase in variance among synthetics as their size is reduced, but a dcerease in their overall mean (Gallais, 1967, Busbice, 1969, Hill, 1971). As a rule of thumb, a compromise lying between 4 and 10 has been recommended (Kinman and Sprague, 1945, Gallais, Guy, and Lenoble 1970). However, the identity and number of parents in the best possible synthetic which can be produced from a given population will depend on the exact distribution of genotypes at each locus, and be a specific property of the population. A method of direct estimation of the expected performance of any synthetic would allow this difficulty to be overcome.

Literature

Busbice, T. H.: Inbreeding in synthetic varieties. Crop Sci. **9**, 601–604 (1969).

Comstock, R. E., Robinson, H. F., Harvey, P. H.: A breeding procedure designed to make maximum use of both general and specific combining ability. Agron. J. **41**, 360–367 (1949).

Falconer, D. S.: Introduction to quantitative genetics. Edinburgh: Oliver and Boyd 1960.

Fisher, R. A.: The genetic basis of natural selection. Oxford: Clarendon Press 1930.

Gallais, A.: Evolution de la vigeur des variétés synthetiques diploids au cours des générations de multiplication. I. En panmixie, influence du nombre de parents et du coéfficient de consanguinité de départ. Ann. Amélior. Plantes **17**, 291–301 (1967).

Gallais, A., Guy, P., Lenoble, M.: Models for varieties in crossfertilised forage plants. Proc. XI Intern. Grassl. Congr. 254–259 (1970).

Graumann, H. O.: The polycross method of breeding in relation to synthetic varieties and recurrent selection of new clones. Proc. VI Intern. Grassl. Congr. 314–319 (1952).

Griffing, B.: Prediction formulae for general combining ability selection methods utilising one or two randommating populations. Aust. J. Biol. Sci. **15**, 650–665 (1962).

Hill, R. R.: Designs to estimate effects of clone substitution in alfalfa synthetics. Crop Sci. **6**, 471–473 (1966).

Hill, R. R.: The effect of the number of parents on the mean and variance of synthetic varieties. Crop Sci. **11**, 283–286 (1971).

Kinman, M. L., Sprague, G. F.: Relation between number of parental lines and theoretical performance of synthetic varieties of corn. J. Amer. Soc. Agron. **37**, 341–351 (1945).

Wright, A. J.: The selection of parents for synthetic varieties of outbreeding diploid crops. Theor. and Appl. Gen. **43**, 79–82 (1973).

Received May 7, 1974
Communicated by R. Riley

Dr. A. J. Wright
Plant Breeding Institute
Maris Lane
Trumpington
Cambridge CB2 2LQ (England)