THE MODIFIED DIALLEL TABLE WITH PARTIAL INBREEDING AND INTERACTIONS WITH ENVIRONMENT¹

D. F. MATZINGER AND O. KEMPTHORNE

Iowa State College, Ames

Received June 22, 1956

THE reasons for considering the covariances of relatives are firstly that they will be of interest *per se* in indicating the regression of offspring on parent etc., and secondly that they lead to the estimation of genetic parameters which are useful as a description of a genetic population and in the formulation of breeding procedures.

Various methods of obtaining theoretical covariances between relatives have been developed, each with a specific set of assumptions. The problem was first considered for random mating populations in equilibrium by FISHER (1918). Correlations were obtained for parent-offspring, full-sib, half-sib, uncle-nephew, cousin, and double-first-cousin relationships for any degree of dominance. The general solution included an arbitrary number of alleles with two-factor epistacy considered for parent-offspring and full-sib correlations. WRIGHT (1921) proposed the method of path coefficients for the study of correlations between relatives with respect to genes which are strictly additive in their effects.

The analysis of diallel crosses presented by HAYMAN (1954a, 1954b) considered an arbitrary degree of dominance for the case where the parents in the diallel table are completely homozygous. Only two alleles per locus were considered and the assumption of no epistacy was imposed. JINKS (1954) applied this method to a study of height, flowering time, and leaf length in *Nicotiana rustica* to estimate the average degree of dominance. The regression of array covariance on variance was plotted to obtain evidence of non-additive gene effects. Deviations from unit slope indicated non-allelic gene interaction might be responsible for the observed heterosis in some F₁ hybrids. Further applications of this method to actual data were given by JINKS (1955). DICKINSON and JINKS (1956) extended the methods of HAYMAN and JINKS to include inbreeding of the parental material in the diallel table.

COCKERHAM (1954) has considered the covariances between relatives under random mating of an inbred population for the case of two alleles per locus. These covariances are expressible in terms of the components of genotypic variance of the random mating population which was the origin of the inbred population.

GRIFFING (1956) considered diallel crosses of homozygous individuals obtained from a random mating population and presented the genetic parameters in terms of additive, dominance, and epistatic effects for an arbitrary number of loci with arbitrary alleles. The term "modified diallel" was suggested to designate the diallel table where the parents are not included. The variances of specific and general combining ability were given in terms of additive, dominance, and epistatic variances.

ALLARD (1956) has given some procedures for evaluating genotype-environment

¹Contribution from the Departments of Agronomy and Statistics, Journal Paper No. J-3053 of the Iowa Agricultural Experiment Station, Ames, Iowa. Projects 1140 and 890.

interactions in the diallel table. His procedures are a development of the work by HAYMAN and JINKS.

KEMPTHORNE (1956a) removed the assumptions of no epistacy and of only two alleles per locus as given by HAYMAN (1954b) and presented the analysis of diallel crosses for arbitrary alleles, arbitrary loci, and arbitrary epistacy. The parents in the diallel table were homozygous lines obtained by inbreeding without selection from a random mating population. Estimates of the mean and variance of both the random mating and homozygous populations, and the parent-offspring covariance were obtained from the variances and covariances of the diallel table. Evidence was given to show that previous methods of analysis of diallel crosses are of questionable validity.

In the present discussion the previous methods will be extended to include the case of arbitrary inbreeding for the parents in the modified diallel table and the analysis of variance is presented for obtaining estimates of the interactions of genotypic components of variance with environments as represented by locations and years.

COVARIANCES BETWEEN RELATIVES

As a method of determining what happens to the additive, dominance, and epistatic deviations under a random mating system when individuals are inbred and then crossed in all possible crosses, the covariance of full-sibs, $Cov(FS)$, and the covariance of half-sibs, $Cov(HS)$, must be obtained for an arbitrary degree of inbreeding assuming any number of loci with an arbitrary number of alleles.

If two individuals are mated which are more closely related than the average of the population, the resulting progeny are inbred. The two genes at a locus of an individual can be alike because they descend from one gene or because of random sampling from a population. The probability of two genes being identical by descent is F , the inbreeding coefficient, which varies in the range from zero in a random mating population to one in an inbred population. Inbreeding and gene frequency are completely independent, with inbreeding changing the association of the alleles in pairs while gene frequency is the proportion of each allele in the population. Under inbreeding the genotypic array for a single locus with arbitrary alleles is

$$F \sum_i p_i A_i A_i + (1 - F) \sum_{ij} p_i p_j A_i A_j.$$

The genotypic value of an individual y_{ijkl} considering two loci can be written as

$$\begin{aligned} y_{ijkl} = & \mu + \alpha_i^1 + \alpha_j^1 + (\alpha^1 \alpha^1)_{ij} + \alpha_k^2 + \alpha_l^2 \\ & + (\alpha^2 \alpha^2)_{kl} + (\alpha^1 \alpha^2)_{ik} + (\alpha^1 \alpha^2)_{jk} + (\alpha^1 \alpha^2)_{il} \\ & + (\alpha^1 \alpha^2)_{jl} + (\alpha^1 \alpha^2 \alpha^2)_{ikl} + (\alpha^1 \alpha^2 \alpha^2)_{jkl} \\ & + (\alpha^1 \alpha^1 \alpha^2)_{ijk} + (\alpha^1 \alpha^1 \alpha^2)_{ijl} + (\alpha^1 \alpha^1 \alpha^2 \alpha^2)_{ijkl} \end{aligned} \quad (1)$$

where the superscripts indicate the locus (KEMPTHORNE 1956b, 1956c), and for example, α_i^1 is the average effect of the i th allele at locus 1, $(\alpha^2 \alpha^2)_{kl}$ is the dominance effect at locus 2, $(\alpha^1 \alpha^2)_{ik}$ is the interaction of the additive effect of the i th allele at

locus 1 with the additive effect of the k th allele at locus 2, and the other terms have similar meaning.

The detailed derivation of the covariances was given by MATZINGER (1956). Since it is quite lengthy and tedious it will be indicated only for the case of two loci, and the general results will be presented. Linkage is assumed to be absent.

Variance between family means

Consider a cross of $A_i A_j B_k B_l \times A_m A_n B_r B_s$ where A and B are two independent loci, and the subscripts indicate the allele at that locus. Let the frequencies of $A_i A_j$, $B_k B_l$, $A_m A_n$, and $B_r B_s$ be designated as P_{ij} , P'_{kl} , P_{mn} , and P'_{rs} respectively, where for example

$$P_{ij} = (1 - F)p_i p_j + \delta_{ij} F p_i$$

and

$$\begin{aligned} \delta_{ij} &= 1 \text{ if } i = j \\ &= 0 \text{ if } i \neq j. \end{aligned}$$

The variance between family means is obtained as

$$\sum_{\substack{ijkl \\ mnrs}} P_{ij} P'_{kl} P_{mn} P'_{rs} \left[\left\{ \frac{1}{2}(A_i + A_j) \frac{1}{2}(A_m + A_n) \right\} \left\{ \frac{1}{2}(B_k + B_l) \frac{1}{2}(B_r + B_s) \right\} \right]^2$$

in which each genotype symbol is replaced by the corresponding genotypic value and μ is replaced by zero. To solve this expression the terms are written in the form of additive, dominance, and interaction effects of the type in equation (1).

After obtaining the simplification of this equation into components of genotypic variance as given by KEMPTHORNE (1954, 1955) and COCKERHAM (1954), the variance between family means can be written down and extended to the case of any number of loci. In the present case this variance is the covariance of full-sibs and is given by

$$\begin{aligned} Cov(FS) &= \left(\frac{1+F}{2} \right) \sigma_A^2 + \left(\frac{1+F}{2} \right)^2 \sigma_D^2 + \left(\frac{1+F}{2} \right)^2 \sigma_{AA}^2 \\ &+ \left(\frac{1+F}{2} \right)^3 \sigma_{AD}^2 + \left(\frac{1+F}{2} \right)^4 \sigma_{DD}^2 + \left(\frac{1+F}{2} \right)^3 \sigma_{AAA}^2 \\ &+ \left(\frac{1+F}{2} \right)^4 \sigma_{AAD}^2 + \left(\frac{1+F}{2} \right)^5 \sigma_{ADD}^2 + \left(\frac{1+F}{2} \right)^6 \sigma_{DDD}^2 \\ &+ \text{etc.} \end{aligned} \quad (2)$$

Variance between sire means

Let the genotype of an arbitrary sire be $A_i A_j B_k B_l$ with frequency $P_{ij} P'_{kl}$. The dam gametic array is

$$\left\{ \sum_r p_r A_r \right\} \left\{ \sum_s p'_s B_s \right\}$$

and the sire gametic array is

$$\{\frac{1}{2}A_i + \frac{1}{2}A_j\}\{\frac{1}{2}B_k + \frac{1}{2}B_l\}.$$

The offspring array is therefore

$$\{\frac{1}{2}\sum_r p_r A_r A_i + \frac{1}{2}\sum_r p_r A_r A_j\}\{\frac{1}{2}\sum_s p'_s B_s B_k + \frac{1}{2}\sum_s p'_s B_s B_l\}.$$

Substitution in terms of the genotypic values of formula (1) and simplifying leads to the variance between sires for two loci which can easily be extended to arbitrary loci.

Since each sire is mated to dams obtained at random from a random mating population the variance of sire means is the covariance of half-sibs which is given by the equation

$$Cov(HS) = \left(\frac{1+F}{4}\right)\sigma_A^2 + \left(\frac{1+F}{4}\right)^2\sigma_{AA}^2 + \left(\frac{1+F}{4}\right)^3\sigma_{AAA}^2 + \text{etc.} \quad (3)$$

These values for $Cov(FS)$ and $Cov(HS)$ are equivalent to those obtained by COCK-ERHAM (1954) for the case of only two alleles per locus.

STRUCTURE OF MODIFIED DIALLEL TABLE

The experiment to be considered involves all possible crosses between p parents with a given degree of inbreeding, or if one assumes the absence of maternal effects, reciprocal crosses need not be obtained and one has $p(p-1)/2$ families. The analysis discussed below assumes there are no maternal effects.

The analysis of variance of diallel crosses was presented by YATES (1947) and in various other forms by other workers. A model used to analyze the variation of a single experiment is

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + r_k + e_{ijk} \quad (4)$$

where

$$i, j = 1, 2, \dots, p \quad i < j$$

$$k = 1, 2, \dots, r$$

and y_{ijk} is the yield resulting from a cross of the i th line with the j th line grown in the k th replicate; μ is a common mean for all crosses in all replicates; g_i is a measure of the average effect of the i th line to all progeny; s_{ij} is a measure of the average effect of a cross of the i th line and the j th line; r_k is the average effect of all crosses grown in the k th replicate; and e_{ijk} is the experimental error associated with the y_{ijk} observation. The g_i components measure general combining ability, while the s_{ij} 's measure deviations from an additive genetic model, frequently referred to as specific combining ability. All variable components in the model (4) are assumed to have zero expectation, to be uncorrelated, and to have their particular variances.

ROJAS (1951) considered the analysis of variance for tests of hypotheses and estimation of variance components by least squares estimates. The expected mean squares were obtained by finding the expected value of the sums of squares in terms of the original model (4) and dividing by the degrees of freedom.

TABLE 1
Analysis of variance and expected mean squares of modified diallel crosses

Source	d.f.	S.S.	M.S.	E.M.S.*			
				σ^2	σ_s^2	σ_d^2	$\frac{1}{(r-1)} \sum_k r_k^2$
Replicates	$r-1$	$R = \sum_k \frac{Y_{..k}^2}{n} - C$	R'	1	0	0	n
g_i 's	$p-1$	$G = \frac{1}{r(p-2)} \sum_i Y_{i..}^2 - \frac{2(p-1)C}{(p-2)}$	G'	1	r	$r(p-2)$	0
s_{ij} 's	$p(p-3)/2$	$S = \sum_{i,j} \frac{Y_{ij.}^2}{r} - C - G$	S'	1	r	0	0
Error	$(r-1)(n-1)$	$E = T - S - G - R$	E'	1	0	0	0
Total	$rn-1$	$T = \sum_{ijk} y_{ijk}^2 - C$					

* Expectations obtained by ROJAS (1951) above; expectations in terms of concepts of present paper below.

The expected mean squares obtained by ROJAS in relation to variances of general and specific combining ability are presented in table 1. These variances were obtained for the original definitions of general and specific combining ability given by SPRAGUE and TATUM (1942). In the present study the expected mean squares are derived in terms of variances and covariances of relatives with respect to a random mating population by again taking the expectation of the sums of squares in table 1.

Consider a diallel table where a plot is the progeny of a mating of two parents with a known amount of inbreeding. Variances and covariances of plot totals are obtained for a constant number, f , plants per plot. The case of unequal numbers of plants per plot presents a slight complication and is given by KEMPTHORNE (1956c).

All of the individuals in a plot have the same sire and dam so the plot total contains f plants which are full-sibs. The expected value of a plot total is $f\mu_G$ where μ_G is the mean of the genotypic portion of the individuals. Then

$$E(\text{plot total})^2 = f^2\mu_G^2 + f\sigma_G^2 + f(f-1)\text{Cov}(FS)$$

where σ_G^2 is the genotypic variance. The genotypic variance of a plot total is obtained by the usual method as

$$V(\text{plot total}) = E(\text{plot total})^2 - \{E(\text{plot total})\}^2 = f\sigma_G^2 + f(f-1)\text{Cov}(FS) = \sigma_1^2. \quad (5)$$

It is assumed that environment within the experiment is additive with variance σ_e^2 , so that the total variance of a plot total equals $\sigma_e^2 + \sigma_1^2$.

Next consider the covariance of two plots, both with the same sire and dam, but in different replicates. If the parents are not homozygous the individuals in one replicate are not a true duplication of the genotypic value of individuals in the same entry in another replicate. In fact, the individuals are full-sibs. With f plants per plot, the covariance between two plots with the same parents, in different replicates, is

$$f^2 \text{Cov}(FS) = \sigma_2^2. \quad (6)$$

The covariance between plots with one parent in common is a function of the half-sib covariance. With f individuals per plot this covariance will be

$$f^2 \text{Cov}(HS) = \sigma_3^2. \quad (7)$$

The expected mean squares are derived in terms of σ_1^2 , σ_2^2 , and σ_3^2 . The mean is coded as zero so it will not enter into the derivation. Replicates are orthogonal to crosses and their effects can be ignored in most of the computations which follow. Terms involving both the mean and replications will cancel out if carried along. The experimental error will contribute σ_e^2 to each expected mean square and is omitted in the following derivations.

The $Y_{i..}$ term, a sire total, contains $r(p-1)$ plots, and $Y_{i..}^2$ will involve $r(p-1)\sigma_1^2$ terms. Each of the $(p-1)$ plot totals is a full-sib to $r(r-1)$ plot totals in other replicates. Each of $(p-1)$ plot totals in one replication is a half-sib with $(p-2)$ plots in r replicates, over the total of r replicates. Then

$$E(Y_{i..}^2) = r(p-1)\sigma_1^2 + r(r-1)(p-1)\sigma_2^2 + r^2(p-1)(p-2)\sigma_3^2.$$

The grand total, $Y_{...}$, contains a total of rn plots. The square of this term will contain rn terms of σ_1^2 . The full-sib comparisons will be $p/2$ times those for $Y_{i..}$ since all of the dams are also included. Each of n plot totals in one replicate is a half-sib of $2(p-2)$ plots over r replicates, and

$$E(Y_{...}^2) = rn\sigma_1^2 + r(r-1)n\sigma_2^2 + 2r^2(p-2)n\sigma_3^2.$$

Each $Y_{ij.}$ term contains r plots, so $Y_{ij.}^2$ has r terms of σ_1^2 . The full-sib covariances will enter between the replicates $r(r-1)$ times, and there are no half-sib covariances. Then

$$E(Y_{ij.}^2) = r\sigma_1^2 + r(r-1)\sigma_2^2.$$

A total of n plots are contained in each $Y_{..k}$ term. In $Y_{..k}^2$ there will be n terms of σ_1^2 , and no terms in σ_2^2 because there are no comparisons between replicates. Each of n plot totals will involve half-sib covariances with $2(p-2)$ plots resulting in

$$E(Y_{..k}^2) = n\sigma_1^2 + 2n(p-2)\sigma_3^2 + n^2r_k^2.$$

The genotypic part of the expected values of the sums of squares in table 1 are as follows:

$$E(C) = \sigma_1^2 + (r-1)\sigma_2^2 + 2r(p-2)\sigma_3^2$$

$$E(R) = (r-1)\sigma_1^2 - (r-1)\sigma_2^2 + n\sum_k r_k^2$$

$$E(G) = (p-1)\sigma_1^2 + (p-1)(r-1)\sigma_2^2 + r(p-1)(p-4)\sigma_3^2$$

$$E(S) = \frac{1}{2}p(p-3)\sigma_1^2 + \frac{1}{2}(r-1)p(p-3)\sigma_2^2 - rp(p-3)\sigma_3^2$$

$$E(E) = (r-1)(n-1)\sigma_1^2 - (r-1)(n-1)\sigma_2^2$$

$$E(T) = (rn-1)\sigma_1^2 - (r-1)\sigma_2^2 - 2r(p-2)\sigma_3^2.$$

Dividing the expected sums of squares by degrees of freedom gives the expected mean squares in table 1 after slight rearrangement. Each component of variation also contains a term for plot error (σ_e^2) as indicated in table 1.

The similarity between the expected mean squares of ROJAS and those obtained in this paper is very evident in table 1. After substitution for σ_1^2 , σ_2^2 , and σ_3^2 from equations (5), (6), and (7) the equality is

$$\sigma^2 = \sigma_e^2 + f\{\sigma_g^2 - \text{Cov}(FS)\} \quad (8)$$

$$\sigma_s^2 = f^2\{\text{Cov}(FS) - 2\text{Cov}(HS)\}$$

$$\sigma_g^2 = f^2 \text{Cov}(HS). \quad (9)$$

Estimates of these variance components are obtained by equating the expected mean squares to the corresponding observed mean squares and solving for the desired component as follows:

$$\widehat{\sigma_g^2} = \frac{G' - S'}{r(p-2)}$$

$$\widehat{\sigma_s^2} = \frac{S' - E'}{r}$$

$$\widehat{\sigma^2} = E'.$$

$\text{Cov}(FS)$ and $\text{Cov}(HS)$ can be obtained from these estimates as

$$\text{Cov}(\widehat{HS}) = \frac{\widehat{\sigma_g^2}}{f^2}$$

$$\text{Cov}(\widehat{FS}) = \frac{\widehat{\sigma_s^2} + 2\widehat{\sigma_g^2}}{f^2}.$$

From the theoretical derivation of $\text{Cov}(FS)$ and $\text{Cov}(HS)$ in equations (2) and (3), estimates of the components of genotypic variance may be obtained. The values of σ_g^2 and σ_s^2 may be expressed in terms of additive, dominance, and epistatic variances to observe the type of epistatic deviations which contribute to each of these variances. The estimates obtained from equations (8) and (9) and equations (1) and (2) are

$$\begin{aligned} \sigma_g^2 &= f^2 \left\{ \left(\frac{1+F}{4} \right) \sigma_A^2 + \left(\frac{1+F}{4} \right)^2 \sigma_{AA}^2 + \text{etc.} \right\} \\ \sigma_s^2 &= f^2 \left\{ \left(\frac{1+F}{2} \right)^2 \sigma_D^2 + \frac{(1+F)^2}{8} \sigma_{AAA}^2 + \left(\frac{1+F}{2} \right)^3 \sigma_{ADD}^2 \right. \\ &\quad \left. + \left(\frac{1+F}{2} \right)^4 \sigma_{DDD}^2 + \text{etc.} \right\}. \end{aligned}$$

Estimates of σ_A^2 , σ_D^2 , and the epistatic components can be obtained if a series of diallel cross experiments is made with different levels of inbreeding. From the diallel crosses of parents with a single level of inbreeding, estimates of σ_A^2 and σ_D^2 can be

obtained only under the assumption of no epistacy. Estimates of σ_A^2 , σ_D^2 , and σ_G^2 are then obtained as

$$\widehat{\sigma_A^2} = \frac{4\widehat{Cov}(HS)}{1 + F}$$

$$\widehat{\sigma_D^2} = \frac{4\{\widehat{Cov}(FS) - 2\widehat{Cov}(HS)\}}{(1 + F)^2}$$

and

$$\widehat{\sigma_G^2} = \widehat{\sigma_A^2} + \widehat{\sigma_D^2}.$$

For the case of $F = 1$, GRIFFING (1956) showed that the following relationship existed with σ_g^2 and σ_s^2 defined in terms of plot means:

$$\sigma_G^2 = 2\sigma_g^2 + \sigma_s^2.$$

It can be seen that this result is applicable only to the case of $F = 1$. From the expectations in the present paper the general relationship would be

$$2\sigma_g^2 + \sigma_s^2 = Cov(FS)$$

and only for the specific case of $F = 1$ would it be the case that

$$Cov(FS) = \sigma_G^2.$$

The value of F to be used in the formulae depends upon the source of the material used as parents in the diallel table. Normally diallel crosses are made between families and not single individuals. Families which arise from inbreeding without selection in a random mating population in equilibrium will reconstitute the original random mating population when crossed, because the gametic array, and hence the gene frequency, will not change with different levels of inbreeding. The appropriate procedure is to obtain the inbred population by whatever means are most suitable and then to choose individuals at random from this population to serve as parents. Each parent is then selfed to give a sufficient number of plants to enable the crosses to be made. The appropriate F value is that holding before the aforementioned multiplication. In the case of inbred lines in animals one could obtain sufficient individuals by full-sibbing.

ANALYSIS OF EXPERIMENTS COMBINED OVER LOCATIONS AND YEARS

Of major interest in assessing the role of the different types of gene effect and in a practical breeding program is the estimation of the components of genotypic variance over locations and years, and the components of variance due to possible interactions of genotypes with each. An extension of the model for a single experiment to include locations and years is

$$\begin{aligned} y_{ijklt} = & \mu + g_i + g_j + s_{ij} + d_k + y_t + (gd)_{ik} + (gd)_{jk} \\ & + (sd)_{ijk} + (gy)_{it} + (gy)_{jt} + (sy)_{ijt} \\ & + (dy)_{kt} + (gdy)_{ikt} + (gdy)_{jkt} + (sdy)_{ijkt} \\ & + r_{klt} + e_{ijklt} \end{aligned}$$

TABLE 2

*Expected mean squares for components of genotypic variance in experiments
repeated over locations and years*

Source	E.M.S.*
g_i 's	$\sigma_4^2 + r\sigma_{bdy}^2 + br\sigma_{by}^2 + cr\sigma_{bd}^2 + bcr\sigma_5^2 + r(p-2)\sigma_{3dy}^2$ $+ br(p-2)\sigma_{3y}^2 + cr(p-2)\sigma_{3d}^2 + bcr(p-2)\sigma_3^2$
s_{ij} 's	$\sigma_4^2 + r\sigma_{bdy}^2 + br\sigma_{by}^2 + cr\sigma_{bd}^2 + bcr\sigma_5^2$
$(gd)_{ik}$'s	$\sigma_4^2 + r\sigma_{bdy}^2 + cr\sigma_{bd}^2 + r(p-2)\sigma_{3dy}^2 + cr(p-2)\sigma_{3d}^2$
$(sd)_{ijk}$'s	$\sigma_4^2 + r\sigma_{bdy}^2 + cr\sigma_{bd}^2$
$(gy)_{it}$'s	$\sigma_4^2 + r\sigma_{bdy}^2 + br\sigma_{by}^2 + r(p-2)\sigma_{3dy}^2 + br(p-2)\sigma_{3y}^2$
$(sy)_{ijt}$'s	$\sigma_4^2 + r\sigma_{bdy}^2 + br\sigma_{by}^2$
$(gdy)_{ikt}$'s	$\sigma_4^2 + r\sigma_{bdy}^2 + r(p-2)\sigma_{3dy}^2$
$(sdy)_{ijkt}$'s	$\sigma_4^2 + r\sigma_{bdy}^2$
Error	σ_4^2

* Expected mean squares coded as

$$\begin{aligned}\sigma_4^2 &= \sigma_e^2 + \sigma_1^2 - \sigma_2^2 \\ \sigma_{bdy}^2 &= \sigma_{2dy}^2 - 2\sigma_{3dy}^2 \\ \sigma_{by}^2 &= \sigma_{2y}^2 - 2\sigma_{3y}^2 \\ \sigma_{bd}^2 &= \sigma_{2d}^2 - 2\sigma_{3d}^2 \\ \sigma_5^2 &= \sigma_2^2 - 2\sigma_3^2\end{aligned}$$

where y_{ijkltq} is the yield of a cross of the i th and j th lines in the q th replicate in the k th location in the t th year; μ is a common mean for all crosses in r replicates, b locations, and c years; g_i is a measure of the average effect of the i th line to all progeny; s_{ij} is a measure of the average effect of a cross of the i th line with the j th line; d_k is the average effect of the k th location; y_t is the average effect of the t th year; $(gd)_{ik}$ is the average interaction of the i th line with the k th location; and the other interactions have the appropriate meaning designated by the subscripts. In all cases the averages are over the populations of other factors. All parameters in the model except μ and r_{ktq} are considered random with expectation zero and zero correlation.

The components of variance of general and specific combining ability over locations and years were presented by ROJAS and SPRAGUE (1952). The expected mean squares in terms of genotypic components can be extended over locations and years, because crosses and hence comparisons of the g_i 's or s_{ij} 's are orthogonal to locations and years. These are presented in table 2.

The values of σ_1^2 , σ_2^2 , and σ_3^2 have the same meanings as in the individual experiments except that they are defined as variances or covariances of averages over the population of environments which is assumed to be indefinitely large. The addition of d or y to these terms indicates the corresponding interactions. For example, σ_{2dy}^2 is the 3-way interaction of locations and years with a pseudo-factor x , say, which gives rise to the variance σ_2^2 .

Because of the relationships between the expected mean squares in table 2 and those of ROJAS and SPRAGUE, estimates of $Cov(HS)$, $Cov(FS)$, and all interaction terms can be obtained from the following relationships:

$$\begin{aligned}
\sigma_g^2 &= \sigma_3^2 = f^2 \text{Cov}(HS) \\
\sigma_s^2 &= \sigma_6^2 = f^2 \{ \text{Cov}(FS) - 2\text{Cov}(HS) \} \\
\sigma_{gd}^2 &= \sigma_{3d}^2 = f^2 \text{Cov}(HS)_d \\
\sigma_{sd}^2 &= \sigma_{6d}^2 = f^2 \{ \text{Cov}(FS)_d - 2\text{Cov}(HS)_d \} \\
\sigma_{gy}^2 &= \sigma_{3y}^2 = f^2 \text{Cov}(HS)_y \\
\sigma_{sy}^2 &= \sigma_{6y}^2 = f^2 \{ \text{Cov}(FS)_y - 2\text{Cov}(HS)_y \} \\
\sigma_{gdy}^2 &= \sigma_{3dy}^2 = f^2 \text{Cov}(HS)_{dy} \\
\sigma_{sdy}^2 &= \sigma_{6dy}^2 = f^2 \{ \text{Cov}(FS)_{dy} - 2\text{Cov}(HS)_{dy} \} \\
\sigma^2 &= \sigma_4^2 = \sigma_e^2 + f \{ \sigma_g^2 - \text{Cov}(FS) \}
\end{aligned}$$

The term $\text{Cov}(HS)_d$, for example, is the component for the interaction of the likeness of half-sibs with locations. Estimates of the components of variance are obtained by equating the expected mean squares to the observed mean squares and solving for the desired components.

Obtaining estimates of additive, dominance, and epistatic components and interactions with locations and years depends upon the number of levels of inbreeding. For example, for modified diallel crosses of one level of inbreeding in experiments repeated over locations and years only additive and dominance variances and corresponding interactions can be estimated assuming no epistacy by the following equations:

$$\begin{aligned}
\sigma_A^2 &= \frac{4\text{Cov}(HS)}{(1+F)} \\
\sigma_D^2 &= \frac{4\{ \text{Cov}(FS) - 2\text{Cov}(HS) \}}{(1+F)^2} \\
\sigma_{Ay}^2 &= \frac{4\text{Cov}(HS)_y}{(1+F)} \\
\sigma_{Dy}^2 &= \frac{4\{ \text{Cov}(FS)_y - 2\text{Cov}(HS)_y \}}{(1+F)^2} \\
\sigma_{Ad}^2 &= \frac{4\text{Cov}(HS)_d}{(1+F)} \\
\sigma_{Dd}^2 &= \frac{4\{ \text{Cov}(FS)_d - 2\text{Cov}(HS)_d \}}{(1+F)^2} \\
\sigma_{A dy}^2 &= \frac{4\text{Cov}(HS)_{dy}}{(1+F)} \\
\sigma_{D dy}^2 &= \frac{4\{ \text{Cov}(FS)_{dy} - 2\text{Cov}(HS)_{dy} \}}{(1+F)^2} \\
\sigma_G^2 &= \sigma_A^2 + \sigma_D^2 \\
\sigma_{Gy}^2 &= \sigma_{Ay}^2 + \sigma_{Dy}^2 \\
\sigma_{Gd}^2 &= \sigma_{Ad}^2 + \sigma_{Dd}^2 \\
\sigma_{G dy}^2 &= \sigma_{A dy}^2 + \sigma_{D dy}^2
\end{aligned}$$

If the original population had two equally frequent alleles at each variable locus, one could estimate the average squared degree of dominance by $2\hat{\sigma}_D^2/\hat{\sigma}_A^2$.

EXTENSION TO OTHER EXPERIMENTS

With more complete data estimates can be obtained on epistatic parameters and all possible interactions with environment. Estimates from diallel crosses within a given level of inbreeding of the parents can be combined to give information on these components.

The present discussion has been for the analysis of plot totals. Additional information can be obtained from the modified diallel table if individual plant data are obtained. The procedure then is an analysis of variance between and within plots. The expectation of the mean square within plots is

$$\sigma_p^2 + \{\sigma_g^2 - Cov(FS)\}$$

where σ_p^2 is the error attributed to each individual within the plot.

LINKAGE BIAS

In this discussion the genotypic frequencies were assumed to be in linkage equilibrium. However, even if the population is in linkage equilibrium these linkages will affect the covariances, as shown by COCKERHAM (1956) for the case when F equals zero. Assuming no position effects, the covariances between relatives where one is an ancestor of the other are not affected by linkage. However, for full-sibs and half-sibs, linkage causes a bias in the epistatic components of variance, but not in the additive or dominance components. A formulation of the general effects of linkage is one of the outstanding problems in the whole area of quantitative inheritance.

SUMMARY

The theoretical covariances between full-sibs and half-sibs under random mating in an arbitrarily inbred population are presented for the case of arbitrary numbers of alleles and loci, and arbitrary epistasy.

The results are applied to the modified diallel cross table in which the parents are of the same known degree of inbreeding. The theoretical relationships between general combining ability and specific combining ability variances, and the components of genotypic variance in the random mating population are obtained.

The procedures for analysis and interpretation of modified diallel crosses replicated over locations and years are presented.

LITERATURE CITED

- ALLARD, R. W., 1956 The analysis of genetic-environmental interactions by means of diallel crosses. *Genetics* **41**: 305-318.
- COCKERHAM, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859-882.
- 1956 Effects of linkage on the covariances between relatives. *Genetics* **41**: 138-141.
- DICKINSON, A. G., and J. L. JINKS, 1956 A generalised analysis of diallel crosses. *Genetics* **41**: 65-78.
- FISHER, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh* **52**: 399-433.

- GRIFFING, B., 1956 A generalised treatment of the use of diallel crosses in quantitative inheritance. *Heredity* **10**: 31-50.
- HAYMAN, B. I., 1954a The analysis of variance of diallel tables. *Biometrics* **10**: 235-244.
- 1954b The theory and analysis of diallel crosses. *Genetics* **39**: 789-809.
- JINKS, J. L., 1954 The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. *Genetics* **39**: 767-788.
- 1955 A survey of the genetical basis of heterosis in a variety of diallel crosses. *Heredity* **9**: 223-238.
- KEMPTHORNE, O., 1954 The correlation between relatives in a random mating population. *Proc. Roy. Soc. London B* **143**: 103-113.
- 1955 The theoretical values of correlations between relatives in random mating populations. *Genetics* **40**: 153-167.
- 1956a The theory of the diallel cross. *Genetics* **41**: 451-459.
- 1956b The correlations between relatives in random mating populations. *Cold Spring Harbor Symposia Quant. Biol.* 20 (in press).
- 1956c *An Introduction to Genetic Statistics*. New York: John Wiley & Sons. (in press).
- MATZINGER, D. F., 1956 Components of variance of diallel crosses of maize in experiments repeated over locations and years. Unpublished Ph.D. Thesis. Iowa State College Library. Ames, Iowa.
- ROJAS, B. A., 1951 Analysis of a group of experiments on combining ability in corn. Unpublished M. S. Thesis. Iowa State College Library. Ames, Iowa.
- ROJAS, B. A., and G. F. SPRAGUE, 1952 A comparison of variance components in corn yield trials. III. General and specific combining ability and their interaction with locations and years. *Agron. J.* **44**: 462-466.
- SPRAGUE, G. F., and L. A. TATUM, 1942 General vs. specific combining ability in single crosses of corn. *J. Am. Soc. Agron.* **34**: 923-932.
- WRIGHT, S., 1921 Systems of mating, I-V. *Genetics* **6**: 111-178.
- YATES, F. 1947 Analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity* **1**: 287-301.