ALTERNATIVE ANALYSES OF THE DIALLEL CROSS *

STANLEY WEARDEN

Department of Genetics, University of Birmingham †

1. INTRODUCTION

A FULL diallel cross consists of the p^2 possible matings among a set of p parental lines including $\frac{1}{2}p(p-1)$ pairs of reciprocal crosses. The purpose of making a diallel cross is to obtain knowledge of the kinds and magnitudes of variability which contribute to differences among the lines. When the geneticist approaches diallel analysis, however, he is often overwhelmed by the many and seemingly different analyses presented in the literature. Yates (1947) has given an analysis; this has been modified and stated in terms of biometrical genetic parameters by Jinks and Hayman (1953) and further by Hayman (1954a, b) and Jinks (1954). Kempthorne (1956) has discussed these methods of analysis in terms of variances of inbred parents, crossbred offspring and the covariance between parents and offspring. Griffing (1958) has classified the four different ways in which at least a $p(p-1)/2$ set of crosses can be obtained, and has discussed the analysis of each in terms of the variances of general and specific combining ability. In order to estimate not only combining ability variances but that of maternal effects as well, Henderson (1948 and 1952) has used a different analysis. Also concerned with maternal effects, Jinks (1954) and Jinks and Broadhurst (1963) have used still another analysis of a diallel cross.

The purpose of this paper is to discuss a diallel cross replicated in a randomised complete block design. Biology dictates that two different models are possible, and statistics define two methods of sampling. Three analyses will be considered under each model and method of sampling, and indications given as to the appropriate analyses for particular biological and statistical situations.

2. STATISTICAL DESIGN AND LINEAR MODELS

(i) Design

As stated in the introduction, it has been assumed that the p^2 possible matings are replicated in a randomised complete block design. This design is not essential for diallel analysis, but replication is generally necessary in order to obtain an estimate of random variation.

The sources of variation, degrees of freedom and expectations of mean squares for the analysis of such a design are

Source	d.f.	E (M.S.)
Replicates. . .	$b-1$	$\sigma^2 + p^2\sigma_B^2$
Matings . . .	p^2-1	$\sigma^2 + b\sigma_M^2$
Random variation .	$(b-1)(p^2-1)$	σ^2

The objective in analysing a diallel cross is the partitioning into meaningful components of the p^2-1 degrees of freedom for matings, the corresponding sums of squares, and the variance among matings (σ_M^2). In order to do this, a linear model based on the biology of the species is required. The model states the nature of the effects producing variability, and the appropriate analysis allows for the estimation of the magnitudes of these different kinds of variability.

(ii) Model for maternal effects (m.m.e.)

In diploid species, the male and female contribute equally to the nuclear genetic composition of the zygote, but their biological contributions are not always equal. Even if cytoplasmic inheritance is ruled out, other maternal effects are possible. The female gamete is often larger than the male; the fruit of plants may contain maternal tissue or endosperm with a greater maternal contribution; many species of animals are viviparous, and others retain or incubate the zygote during a part of its development. Finally, many of the vertebrates supply food to their young after they are born or hatched. Thus a model which allows for differences among maternal effects reflects a common biological situation. The model for maternal effects is

$$Y_{ijk} = \mu + g_i + g_j + m_j + s_{ij} + b_k + \epsilon_{ijk}$$

where

μ = a mean common to all matings to which inference can be made from this p^2 set.

g_i = the common genic contributions of the ith paternal line.

g_j = the common genic contribution of the jth maternal line.

m_j = the maternal contribution of the jth line.

s_{ij} = the interaction between the genetic contribution of the ith line and that of the jth line.

b_k = the effect of the kth replicate; it is commonly assumed that the b_k are normally, independently distributed with mean zero and variance σ_B^2.

ϵ_{ijk} = the random effect peculiar to the ijth cross in replicate k; the ϵ_{ijk} are assumed n.i.d. (o, σ^2).

The nature and distributions of g_i, g_j, m_j and s_{ij} depend on the sampling method and will be discussed in these sections.

(iii) *Model for reciprocal effects (m.r.e.)*

Many examples can be cited where the progeny of the ij cross differ significantly from those of the ji cross, and these differences can not always be assigned to maternal effects. Sex linkage is probably the most simple explanation of such reciprocal effects, but no matter what the cause, it is not an unusual occurrence. A diallel model which includes reciprocal effects is

$$Y_{ijk} = \mu + g_i + g_j + s_{ij} + r_{ij} + b_k + \epsilon_{ijk}$$

where

μ = a mean common to all matings to which inference can be made from this p^2 set.

g_i = the common genic contribution of the ith paternal line.

g_j = the common genic contribution of the jth maternal line.

s_{ij} = the interaction between the genetic contribution of the ith line and that of the jth line.

r_{ij} = the additional effect of using the ith line as male parent and the jth line as female parent; $r_{ij} = -r_{ji}$, and $r_{ii} = 0$.

b_k = the effect of the kth replicate, again the b_k are assumed n.i.d. $(0, \sigma_B^2)$.

ϵ_{ijk} = the random effect peculiar to the ijth cross in replicate k; the ϵ_{ijk} are assumed n.i.d. $(0, \sigma^2)$.

Thus the two models differ only in that one allows for maternal effects and the other for reciprocal effects. It is possible for both effects to appear within a single set of crosses. A model containing both can be written, but as will be seen later, the analysis of other effects becomes more difficult.

(iv) *Combining ability*

The meaning of g_i, g_j and s_{ij} is discussed by Hayman (1954a, b) and Kempthorne (1956) in biometrical genetic terms, and in terms of general and specific combining ability by Griffing (1958). It should be pointed out that g_i is not general combining ability as defined by Sprague and Tatum (1942), nor does s_{ij} fit their definition of specific combining ability, for they defined these effects as deviations from the mean of the F_1's. In this paper, these effects are deviations from the weighted mean of the p inbreds and the $p(p-1)$ crossbreds. If there is any difference between g_i and general combining ability, it is probably trivial; s_{ij}, in addition to Sprague-Tatum specific combining ability, contains a component for the average deviation of the crosses from their respective midparents. This component was termed the " mean dominance deviation " by Hayman (1954a) who presented a method for computing a sum of squares and testing this fraction of s_{ij} variation for significance.

3. SAMPLING METHODS

(i) *Random sample.*

In nature, a species is often subdivided into easily identifiable groups or lines such as clones, varieties, families, herds, flocks, etc. To include all possible subgroups in a single diallel experiment is usually impossible, so a random sample of p lines is used in the experiment and inference extended to the entire population. Such an experiment thus falls in the category of Eisenhart's (1947) Model II, and the expectations of means and variances are accordingly affected.

With random sampling, μ is the population mean, and the expectations of g_i, g_j, s_{ij} and m_j are all zero. The expectations of these terms squared are

$$E(g_i^2) = E(g_j^2) = \sigma_g^2$$
$$E(s_{ij}^2) = \sigma_s^2 \text{ and}$$
$$E(m_j^2) = \sigma_m^2.$$

With respect to reciprocal effects, they are fixed in the sense that $r_{ij} = -r_{ji}$, but the $p(p-1)/2$ pairs of reciprocal effects can be assumed to come from a large population of such pairs of effects, and thus $E(r_{ij}^2) = \sigma_r^2$. It is essential to the analysis to assume that all effects are independently distributed, and in order to make tests of significance, it is necessary to assume they are normally distributed as well.

(ii) *Fixed sample*

The breeder may be interested in the variability among crosses of certain particular lines. These lines are not a random sample, but have been chosen by him because of certain merits possessed by each. Model I of Eisenhart (1947) describes such experiments, and again the expectations are affected accordingly.

For a fixed group of lines, μ is no longer the population mean, but the mean of all possible replications of the experiment. By saying that effects are *fixed* is meant that they sum to zero. Thus, $\Sigma g_i = 0$, $\Sigma g_j = 0$, and $\Sigma m_j = 0$; in addition $\Sigma s_{ij} = 0$ for all i and j. However, it is still assumed that $\Sigma r_{ij} = 0$ only for each ij combination. The effects g_i, g_j, m_j and s_{ij} are constants for the fixed sample, and the expectations of their squares are the constants squared. In the usual sense, they have no variance, but one can compute an average of squared effects; this is done when one computes a *mean square* in the analysis of variance of an experiment containing fixed effects. This average of squared effects is computed by dividing the sum of squared constants by the appropriate degrees of freedom. To avoid superfluous symbols, these averages of squared constants will be called "variances of fixed effects", and the same σ^2 symbolism will be used as for random sampling,

except that the " variances " will be marked with a tilde superscript. Hence, by definition

$$\Sigma g_i^2/p - 1 = \Sigma g_j^2/p - 1 = \tilde{\sigma}_g^2 = p\sigma_g^2/p - 1$$

$$\Sigma s_{ij}^2/(p-1)^2 = \tilde{\sigma}_s^2 = p^2\sigma_s^2/(p-1)^2 \text{ and}$$

$$\Sigma m_j^2/p - 1 = \tilde{\sigma}_m^2 = p\sigma_m^2/p - 1.$$

4. ANALYSES

The three analyses to be considered are (1) that as outlined by Hayman (1954), but without the subdivision of what he termed the b sum of squares, (2) that of a p by p factorial such as used by Jinks and Broadhurst (1963) and (3) that which the author has inferred from the sums of squares given by Henderson (1952). The second is the customary analysis of a two-way table of data; the third is a refinement which subdivides the interaction term of the factorial analysis. Certain terms are common in two different analyses, but for ease of reference the terminology of the original papers has been maintained as much as possible despite the redundance.

To keep mathematical symbolism to a minimum, each analysis will be for a single replicate of a p^2 diallel cross. To expand the analysis to cover an experiment containing b replicates, crosses are summed across all replicates and the analysis performed on these totals. All divisors for sums of squares must be multiplied by b as are the coefficients of all variances except σ^2 in the expectations of mean squares.

5. SUMS OF SQUARES AND EXPECTATIONS

(i) Random sample

In order to subdivide the $p^2 - 1$ degrees of freedom, and the corresponding sums of squares, for matings according to the three analyses to be considered, it is necessary to compute eight sums of squares. Sums of squares for a single replicate are given in table 1 along with their expectations, assuming a random sample of lines. Expectations of σ_m^2 under the model for maternal effects (m.m.e.) and for σ_r^2 under the model for reciprocal effects (m.r.e.) are also given.

In table 1, a dot indicates summation from 1 to p over the values with the omitted subscript, and the sigmas indicate summation over all values of i or ij combinations. Among the letter coded symbols, a prime indicates that the sum of squares expresses variation from the origin (or zero), while letters without the prime indicate that the variability is from the mean. The arabic numeral 1 is used as the symbol for the " correction term " because of the one degree of freedom associated with this value. Numerical examples of the computation of all values except P′ and M′ can be found in the paper by Hayman (1954a), whereas these two values are simply the " uncorrected row "

TABLE 1

Coefficients for variation from origin in a single replicate of a diallel cross among p lines chosen randomly

Sum of squares	Symbol	μ^2	σ_g^2	σ_s^2	σ^2	m.m.e. σ_m^2	m.r.e. σ_r^2
ΣY_{ij}^2	T'	p^2	$2p(p+1)$	p^2	p^2	p^2	$p(p-1)$
$\Sigma (Y_{i\cdot})^2/p$	P'	p^2	$p(p+3)$	p	p	p	$p-1$
$\Sigma (Y_{\cdot j})^2/p$	M'	p^2	$p(p+3)$	p	p	p^2	$p-1$
$Y_{\cdot\cdot}^2/p^2$	1	p^2	$4p$	1	1	p	0
$\Sigma (Y_{i\cdot}+Y_{\cdot i})^2/2p$	G'	$2p^2$	$2p(p+3)$	$2p$	$p+1$	$\dfrac{p}{2}(p+3)$	0
$\underset{i>j}{\Sigma} (Y_{ij}+Y_{ji})^2/2+\Sigma Y_{ii}^2$	C'	p^2	$2p(p+1)$	p^2	$\dfrac{p}{2}(p+1)$	$\dfrac{p}{2}(p+1)$	0
$\Sigma (Y_{i\cdot}-Y_{\cdot i})^2/2p$	R	0	0	0	$p-1$	$\dfrac{p}{2}(p-1)$	$2(p-1)$
$\underset{i>j}{\Sigma} (Y_{ij}-Y_{ji})^2/2$	D	0	0	0	$\dfrac{p}{2}(p-1)$	$\dfrac{p}{2}(p-1)$	$p(p-1)$

TABLE 2

Expectations of mean squares for a single replicate of a p^2 diallel cross when parental lines are random

Source	d.f.	S.S.	m.m.e. E (M.S.)	m.r.e. E (M.S.)
Hayman Analysis				
a (Parental lines)	$p-1$	$G'-2(1)$	$\sigma^2+\dfrac{p}{2}\sigma_m^2+2\sigma_s^2+2p\sigma_g^2$	$\sigma^2+2\sigma_s^2+2p\sigma_g^2$
b (Genetic interaction)	$\dfrac{p}{2}(p-1)$	$C'-G'+1$	$\sigma^2+2\dfrac{(p-1)}{p}\sigma_s^2$	$\sigma^2+2\dfrac{(p-1)}{p}\sigma_s^2$
c (Av. maternal effects)	$p-1$	R	$\sigma^2+\dfrac{p}{2}\sigma_m^2$	$\sigma^2+2\sigma_r^2$
d (Reciprocal effects)	$\dfrac{1}{2}(p-1)(p-2)$	$D-R$	σ^2	$\sigma^2+2\sigma_r^2$
Factorial Analysis				
Maternal strains	$p-1$	$M'-1$	$\sigma^2+\sigma_s^2+p\sigma_g^2+p\sigma_m^2$	$\sigma^2+\sigma_r^2+\sigma_s^2+p\sigma_g^2$
Paternal strains	$p-1$	$P'-1$	$\sigma^2+\sigma_s^2+p\sigma_g^2$	$\sigma^2+\sigma_r^2+\sigma_s^2+p\sigma_g^2$
$M\times P$	$(p-1)^2$	$T'-P'-M'+1$	$\sigma^2+\sigma_s^2$	$\sigma^2+\dfrac{p-2}{p-1}\sigma_r^2+\sigma_s^2$
Henderson Analysis				
Dams	$p-1$	$M'-1$	$\sigma^2+\sigma_s^2+p\sigma_g^2+p\sigma_m^2$	$\sigma^2+\sigma_r^2+\sigma_s^2+p\sigma_g^2$
Sires	$p-1$	$P'-1$	$\sigma^2+\sigma_s^2+p\sigma_g^2$	$\sigma^2+\sigma_r^2+\sigma_s^2+p\sigma_g^2$
Crosses	$\dfrac{p}{2}(p-1)$	$C'-G'+1$	$\sigma^2+2\dfrac{(p-1)}{p}\sigma_s^2$	$\sigma^2+\dfrac{2(p-1)}{p}\sigma_s^2$
Remainder	$\dfrac{1}{2}(p-1)(p-2)$	$D-R$	σ^2	$\sigma^2+2\sigma_r^2$

and " uncorrected column " sums of squares as they are usually computed for a two-way analysis of variance.

Degrees of freedom, methods of computing the sums of squares, and expectations of mean squares under m.m.e. and m.r.e. are given respectively for each method of analysis in table 2.

Under the model for maternal effects, the Hayman analysis yields valid variance ratio tests for the significance of maternal effects in the ratio of *c mean square/d mean square*. The ratio of *b mean square/d mean square* gives a valid variance ratio test for the significance of the variance arising from genetic interaction. However, under this model there is no test for the significance of genic variance. If the underlying biology dictates the model for reciprocal effects, the *c* and *d* mean squares have the same expectations. Under such circumstances, these two terms in the analysis of variance can be pooled, as was done by Griffing (1958). With this model and replication, the significance of reciprocal effects and of genetic interaction variance can be evaluated by testing the pooled *c* and *d* mean square and the *b* mean square, respectively, by the random variation mean square. While there is no exact test of significance for genic variance, there is a conservative variance ratio test. By a conservative test is meant that the actual probability of a type I error is less than the α-level given in a table of critical values. Such a test for genic variance is the ratio *a mean square/b mean square*. Under either model, all variance components can be estimated.

Using the factorial analysis, there are valid variance ratio tests only under the model for maternal effects. Maternal effects are tested for significance by the ratio *Maternal mean square/Paternal mean square*, while genic variance is tested for significance by the ratio *Paternal mean square/MXP mean square*. To make a test of significance for variance due to genetic interaction, replication is necessary, for there must be an estimate of random variation in order to obtain the ratio *M × P mean square/Random variation mean square*. If the model for reciprocal effects is the correct assessment of the biological situation, the factorial analysis is worthless. Not only are there no valid variance ratio tests, but it is not even possible to estimate σ_r^2, σ_s^2 and σ_g^2. Because of the redundance in expectations of the Paternal and Maternal mean squares, there are but two independent equations with which to solve for three unknowns.

The analysis which is inferred from Henderson's equations for estimation of variance components is particularly intended for the model for maternal effects. The significance of maternal effects can be evaluated by testing the Dams mean square by the Sires mean square. The ratio of Crosses mean square to Remainder mean square is an exact test of significance for the variance arising from genetic interaction. In the ratio *Sires mean square/Crosses mean square* there is a conservative test for the significance of genic variance, but it may be too conservative if this type of variance is relatively small, *i.e.* when the trait under study has low *heritability*. Under m.m.e., all variance

components can be estimated. If the model for reciprocal effects describes the biological sources of variation better than that for maternal effects, the Henderson analysis yields valid tests for the significance of σ_r^2 and σ_s^2 if there is replication. To obtain these, Remainder mean square and Crosses mean square are tested, respectively, by Random variation mean square. There is no test for the significance of σ_g^2. Estimation of σ_g^2 σ_s^2 and σ_r^2 is possible; in doing this, the Dams and Sires items of the analysis of variance should be pooled.

(ii) Fixed sample

The sums of squares necessary for the three analyses of a diallel cross involving a fixed set of p lines are given in table 3. This table

TABLE 3

Coefficients for variation from origin in a single replicate of diallel cross among p fixed lines

Sum of squares	Symbol	μ^2	σ_g^2	σ_s^2	σ^2	m.m.e. σ_m^2	m.r.e. σ_r^2
ΣY_{ij}^2	T'	p^2	$2p^2$	p^2	p^2	p^2	$p(p-1)$
$\Sigma(Y_{i\cdot})^2/p$	P'	p^2	p^2	0	p	0	$p-1$
$\Sigma(Y_{\cdot j})^2/p$	M'	p^2	p^2	0	p	p^2	$p-1$
$Y_{\cdot\cdot}^2/p^2$	I	p^2	0	0	1	0	0
$\Sigma(Y_{i\cdot}+Y_{\cdot i})^2/2p$	G'	$2p^2$	$2p^2$	0	$p+1$	$\dfrac{p^2}{2}$	0
$\underset{i>j}{\Sigma}(Y_{ij}+Y_{ji})^2/2+\Sigma Y_{ii}^2$	C'	p^2	$2p^2$	p^2	$\dfrac{p}{2}(p+1)$	$\dfrac{p^2}{2}$	0
$\Sigma(Y_{i\cdot}-Y_{\cdot i})^2/2p$	R	0	0	0	$p-1$	$\dfrac{p^2}{2}$	$2(p-1)$
$\underset{i>j}{\Sigma}(Y_{ij}-Y_{ji})^2/2$	D	0	0	0	$\dfrac{p}{2}(p-1)$	$\dfrac{p^2}{2}$	$p(p-1)$

also includes coefficients for the " variances " of the fixed genic effects and genetic interaction as well as that for maternal effects under m.m.e. and coefficients for σ_r^2 under m.r.e.

The coefficients in table 3 differ in many instances from the corresponding ones for random sampling. There are two reasons for this; the first being the fact that fixed effects, when summed over all values, add to zero. Again because they sum to zero, there is a negative correlation within any set of p constants. Using the terminology of " variances " of fixed effects, the expected covariance between any two constants, say g_i and g_j, in the same set is $-\sigma_g^2/p-1$. Thus,

$$E(g_i+g_j)^2 = 2(\sigma_g^2-\sigma_g^2/p-1).$$

Table 4 contains the expectations of mean squares for the three different analyses of the diallel cross when there is a fixed set of p parental lines. In computing the expectations of mean squares, it is

necessary to remember the mathematical definitions of the " variances " which were given in Section 3(ii).

The Hayman analysis was intended to deal with experiments involving a fixed set of lines. Indeed, the analysis even provides for some g_i and g_j being identical by descent, this situation being detected by Hayman's b_2. Other implications of significance of the various

TABLE 4

Expectations of mean squares for a single replicate of a p^2 diallel cross when parental lines are fixed

Source	d.f.	S.S.	E (M.S.) for m.m.e.	E (M.S.) for m.r.e.
Hayman Analysis				
a (Parental lines)	$p-1$	$G'-2(1)$	$\sigma^2 + \frac{p}{2}\tilde{\sigma}^2_m + 2p\tilde{\sigma}^2_g$	$\sigma^2 + 2p\tilde{\sigma}^2_g$
b (Genetic interaction)	$\frac{p}{2}(p-1)$	$C'-G'+1$	$\sigma^2 + \frac{2p}{p-1}\tilde{\sigma}^2_s$	$\sigma^2 + \frac{2p}{p-1}\tilde{\sigma}^2_s$
c (Av. maternal effects)	$p-1$	R	$\sigma^2 + \frac{p}{2}\tilde{\sigma}^2_m$	$\sigma^2 + 2\sigma^2_r$
d (Reciprocal effects)	$\frac{1}{2}(p-1)(p-2)$	$D-R$	σ^2	$\sigma^2 + 2\sigma^2_r$
Factorial Analysis				
Maternal strains	$p-1$	$M'-1$	$\sigma^2 + p\tilde{\sigma}^2_g + p\tilde{\sigma}^2_m$	$\sigma^2 + \sigma^2_r + p\tilde{\sigma}^2_g$
Paternal strains	$p-1$	$P'-1$	$\sigma^2 + p\tilde{\sigma}^2_g$	$\sigma^2 + \sigma^2_r + p\tilde{\sigma}^2_g$
M×P	$(p-1)^2$	$T'-P'-M'+1$	$\sigma^2 + \tilde{\sigma}^2_s$	$\sigma^2 + \frac{p-2}{p-1}\sigma^2_r + \tilde{\sigma}^2_s$
Henderson Analysis				
Dams	$p-1$	$M'-1$	$\sigma^2 + p\tilde{\sigma}^2_g + p\tilde{\sigma}^2_m$	$\sigma^2 + \sigma^2_r + p\tilde{\sigma}^2_g$
Sires	$p-1$	$P'-1$	$\sigma^2 + p\tilde{\sigma}^2_g$	$\sigma^2 + \sigma^2_r + p\tilde{\sigma}^2_g$
Crosses	$\frac{p}{2}(p-1)$	$C'-G'+1$	$\sigma^2 + \frac{2p}{p-1}\tilde{\sigma}^2_s$	$\sigma^2 + \frac{2p}{p-1}\tilde{\sigma}^2_s$
Remainder	$\frac{1}{2}(p-1)(p-2)$	$D-R$	σ^2	$\sigma^2 + 2\sigma^2_r$

fractions of b have been considered by Jinks (1954). Under m.m.e., the genic variation among parental lines (*a*) can be tested by the mean square dealing with variation due to average maternal effects (*c*); variation due to genetic interaction (*b*) and that due to average maternal effects can each be tested by the *d* mean square, which is an independent estimate of random variation under m.m.e. If there is a possibility of genotype by replicate interaction, Hayman (1954a) has shown how the Random variation sum of squares can be subdivided to provide independent test terms for *b* and *c*. Under m.r.e. all terms can be tested by the Random variation mean square to yield exact variance ratio tests. Again, with suspicion of genotype by replicate

2 X

interaction, the Random variation sum of squares can be subdivided to provide independent test terms. The suitability of the Hayman analysis for a diallel experiment involving a fixed set of parental lines is shown in the fact that there are exact tests of significance for all components of variation under each model, and likewise the magnitude of each component can be estimated.

Under m.m.e., the factorial analysis again yields exact variance ratio tests for all effects when there is replication. The Random variation mean square is used to test the Paternal and the $M \times P$ mean squares for the significance of $\tilde{\sigma}_g^2$ and $\tilde{\sigma}_r^2$, respectively; maternal effects are tested for significance by the ratio *Maternal mean square/Paternal mean square*. The factorial analysis is again completely inappropriate for the model for reciprocal effects, for under this model, the factorial analysis can provide neither tests of significance nor estimates of the magnitudes of the different sorts of variability.

In a diallel experiment involving a species with differential maternal effects, the Henderson analysis permits exact variance ratio tests for all sources of variation when the parental lines are fixed. The significance of maternal effects can be tested by the ratio of Dams mean square to Sires mean square. Both genic effects and genetic interaction can be tested by the respective ratios of Sires mean square and Crosses mean square to the Remainder mean square. Under m.r.e., the Henderson analysis allows for tests of the variation due to genetic interaction and reciprocal effects when there is replication, for the last two terms in the analysis of variance can be tested by Random variation mean square. There is no exact test for genic effects, but a decidedly conservative test does arise from the ratio *Sires mean square/Remainder mean square*. Estimates of the magnitudes of all components of variance can be obtained under either model.

6. REPLICATION AND POWER

As pointed out in a previous section, replication is virtually essential to the analysis of a diallel experiment irrespective of model. In most cases it is needed to provide an estimate of random variation (σ^2). When a comparison within the p^2 matings permits an independent estimate of σ^2, this can be pooled with that resulting from replication to add denominator degrees of freedom, and hence power, to tests of significance in which this source of variation is the test term. In other instances, the existence of genotype by replicate interaction may require that the Random variation sum of squares be subdivided into a number of independent test terms as described by Hayman (1954a) in his numerical example. Corresponding procedures can be employed in the other two analyses.

If one is more interested in assaying the significance of one particular variance, the relative power of different variance ratio tests may dictate the analysis to be used. The following statements on

power assume that there is replication and that all estimates of random variation are pooled. Under both models and both methods of sampling, the equivalent mean squares of the Hayman and Henderson analyses give the more powerful test of σ_s^2. For a fixed set of p lines, the Hayman analysis yields the most powerful test of σ_g^2 under m.r.e., but under m.m.e. the relative power of the Hayman test for σ_g^2 will probably be less than the tests one would employ in the factorial and the Henderson analyses. The presence of a coefficient of σ_m^2 in the expectation of the a mean square under this model necessitates the use of the c mean square as a test term, and there is a consequential loss in denominator degrees of freedom for the variance ratio test. The larger the relative magnitude of σ_m^2 and the greater the loss in degrees of freedom, the more adversely affected will be the power of the test for genic effects within the Hayman analysis under m.m.e. A similar argument can be used to show why the Hayman analysis will probably yield a more powerful test for maternal effects,* irrespective of model. Under m.r.e. the pooling of the c and d mean squares and testing them by the Random variation mean square will yield the most powerful test of σ_r^2, again for both sampling methods.

7. SUMMARY

The expectations of mean square for three analyses of a diallel cross are presented for a model containing maternal effects and for one containing reciprocal effects. The expectations are additionally affected according to whether the p set of parental lines are randomly or selectively chosen.

The appropriate analysis of a given diallel experiment can usually be chosen on the basis of model, the method of selecting parental lines and the relative power of tests of the variances in which the geneticist is particularly interested.

8. REFERENCES

EISENHART, C. 1947. The assumptions underlying the analysis of variance. *Biometrics, 3*, 1-21.

GRIFFING, B. 1958. Application of sampling variables in the identification of methods which yield unbiased estimates of genotypic variance components. *Australian Journal of Biological Sciences, 11*, 219-245.

---

* It was pointed out to the author by J. L. Jinks that the c mean square of the Hayman analysis does not designate whether a significant effect is maternal or paternal in origin. Preparous paternal attention has been described in some species, notably the seahorse, stickleback and catfish; postparous effects due to the male could be expected in patriarchal herds, packs and flocks. Hence, one not completely versed in the biology of the experimental species is advised to use the factorial or the Henderson analysis to designate the sex of parent to which these effects are due.

HAYMAN, B. I. 1954a. The analysis of variance of diallel crosses. *Biometrics, 10,* 235-244.

HAYMAN, B. I. 1954b. The theory and analysis of diallel crosses. *Genetics, 39,* 789-809.

HENDERSON, C. R. 1948. Estimation of general, specific, and maternal combining abilities in crosses among inbred lines of swine. Unpublished *Ph.D. Thesis,* Iowa State College.

HENDERSON, C. R. 1952. Specific and general combining ability. *Heterosis.* Iowa State College Press, Ames.

JINKS, J. L., AND HAYMAN, B. I. 1953. Analysis of diallel crosses. *Maize Genetics Cooperation News Letter, 27,* 48-54.

JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. *Genetics, 39,* 767-788.

JINKS, J. L., AND BROADHURST, P. L. 1963. Diallel analysis of litter size and body weight in rats. *Heredity, 18,* 319-336.

KEMPTHORNE, O. 1956. The theory of the diallel cross. *Genetics, 41,* 451-459.

SPRAGUE, G. F., AND TATUM, L. A. 1942. General vs. specific combining ability in single crosses of corn. *J. Am. Soc. Agron., 34,* 923-932.

YATES, F. 1947. Analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity, 1,* 287-301.