Computational procedure for a weighted diallel analysis*

D. A. Schaff[1]**, G. A. Milliken**[2] **and C. D. Clayberg**[1]

[1] Department of Horticulture and
[2] Department of Statistics; Kansas State University, Manhattan, KS 66506, USA

Summary. A diallel analysis is described for the case in which the values of the characteristic used in an inheritance study have unequal variances. Such a characteristic can be a mean, a slope of a regression line, or the estimates of some parameter of a linear or non-linear model. Computational formulae are presented which incorporate the necessary weighting along with the statistics of the Hayman-Jinks method for diallel analysis. The method described can also be used to perform a weighted diallel analysis based on means when there are unequal numbers of replications per cross. A simple example demonstrates the computations necessary to complete a weighted diallel analysis.

Key words: Heat tolerance − Killing temperature − Inheritance

Introduction

The diallel analysis has been used to study genetic relationships for variables which can be measured directly (e.g., yield, earliness, or flowering) utilizing the mean values of the parents and their crosses. This method may also be used to study the inheritance of phenotypic characteristics which cannot be measured directly, such as the slopes of regression models or other parameters from various models. The estimates of the parameters are obtained by fitting a model to the data from a parent or cross. One problem with the procedure is that the variances of the estimates of the

parameters are not necessarily equal. Inequality of variances can occur when there are unequal numbers of observations per cross, say n_{ij} observations in $cross_{ij}$, and thus the variance of the mean response is $\hat{\sigma}^2/n_{ij}$. A second case occurs when different types of parameters are used in the diallel analysis, such as slopes of regression lines or other parameters from linear or non-linear models used to describe a particular relationship.

The objective of this paper is to modify the equations of Hayman (1954) and Jinks (1954) to conduct a diallel analysis when the estimates of the parameters have unequal variances. The application we use for illustration is a study of the inheritance of heat tolerance as measured by the killing temperature (Schaff 1984).

Materials and methods

To estimate the killing temperature, defined as the temperature at which 50% injury occurs, leaf samples from acclimated bean *Phaseolus vulgaris* L. plants were subjected to heat treatments of 44 °C to 52 °C (Schaff 1984), with incubation after heat treatments changed to 4 h of shaking at 10 °C to reduce the time needed to run the test. The degree of injury, evaluated as electrolyte leakage, was modeled as a function of temperature (Schaff 1984) by the sigmoidal equation:

$$\text{Injury} = 1/(1 + e^{-B(\text{temperature}-U)}) + \varepsilon . \tag{1}$$

The parameter U is the temperature at which 50% injury occurs. The parameters of the model, estimated from the data (Schaff 1984), were the U_{ij}, having variance $\hat{\sigma}^2 K_{ij}^2$. The value of K_{ij} depends on the model and the temperatures at which the data were collected. The estimate of $\hat{\sigma}^2$ is obtained from the plant within

cultivar (ERROR A) mean square (Schaff 1984). Using the estimate of U_{ij}, K_{ij}^2, and $\hat{\sigma}^2$, we can construct a weighted diallel analysis to study the inheritance of heat tolerance as measured by the killing temperature.

Let U_{ij} denote the estimate of the parameter from a model fit to data from the ij^{th} cross. Then the entry in the diallel table for the ij^{th} cross is U_{ij}, where the variance of U_{ij} in general is $\hat{\sigma}^2 K_{ij}^2$. The term K_{ij}^2 is a constant associated with a particular model and depends on the values of independent variables. If U_{ij} is a mean then K_{ij}^2 equals $1/n_{ij}$.

Weighting procedure

The average weight (K_{ij}^2) value is $Koo = \sum_i \sum_j K_{ij}^2/p^2$ where p = number of parents. As in the notation of Hayman (1954) and Jinks (1954), the variances and covariances of the diallel are:

V_r = variance of one array,

W_r = covariance between the parents and their offspring in one array,

V_{ILI} = mean variance of the arrays,

W_{OLOI} = mean covariance between the parents and the arrays,

V_{OLO} = variance of the parents,

V_{OLI} = the variance of the means of the arrays, and

$(M_{LI} - M_{LO})$ = difference between the mean of the parents and the mean of their p^2 progeny.

The following computational equations are needed to evaluate these necessary statistics when the variances of the entries in the diallel table are unequal. The variance of the data from the rth parent's crosses is:

$$V_r = (1/(p-1)) \left[\sum_i (U_{ir}/K_{ir})^2 - \left(U_r^2\left(\sum_i 1/K_{ir}^2\right)\right)\right] Koo,$$

where $U_r = \left(\sum_i U_{ir}/K_{ir}^2\right)\Big/\sum_i 1/K_{ir}^2$, which is the weighted average of the data in column r.

The covariance between the parents and their progeny in the rth column of data is:

$$W_r = (1/(p-1)) \left[\sum_i (U_{ii} - U_p)(U_{ir} - U_r)/K_{ii}K_{ir}\right] Koo.$$

Then $V_{ILI} = 1/p \sum_r V_r$ and $W_{OLOI} = 1/p \sum_r W_r$.

The variance of the parental values is:

$$V_{OLO} = (1/(p-1)) \left[\sum_i (U_{ii}/K_{ii})^2 - \left(U_p^2\left(\sum_i 1/K_{ii}^2\right)\right)\right] Koo,$$

where $U_p = \left(\sum_i U_{ii}/K_{ii}^2\right)\Big/\sum_i 1/K_{ii}^2$, which is the weighted average of the parental values.

Let $S_r = \sum_i 1/K_{ir}^2$. Then the variance of array means is:

$$V_{OLO} = (1/(p(p-1))) \left[\sum_r U_r^2 S_r - \left(\left(\sum_r U_r S_r\right)^2\Big/\sum_r S_r\right)\right] Koo.$$

The square of the difference between the mean of the parents and the mean of their p^2 progeny is:

$$(M_{LI} - M_{LO})^2 = \left[\left(\left(\sum_i \sum_j U_{ij}/K_{ij}^2\right)\Big/\left(\sum_i \sum_j 1/K_{ij}^2\right)\right) - U_p\right]^2.$$

The correlation coefficient between the parents (K_{ii}) and the order of dominance ($W_r + V_r$) is:

$$r = \sum_r \left[\left((W_r + V_r) - \left(\sum_r ((W_r + V_r)/p)\right)\right)(U_{rr} - U_p)/K_{rr}\right]\Big/$$

$$\left(\sum_r (W_r^2 + V_r^2) - \left[\sum_r (W_r + V_r)\right]^2/p\right)$$

$$\cdot\left(\sum_i (U_{ii}/K_{ii})^2 - \left(U_p^2\left(\sum_i 1/K_{ii}^2\right)\right)\right).$$

The error term is multiplied by Koo so that it is in the same units as the above statistics (V_r, W_r, etc.; $\hat{E} = \text{Error A} \times Koo$). The remaining computations in the diallel analysis are unaffected by the unequal variances. The genetic components of the Hayman (1954) and Jinks (1954) model are:

$\hat{D}$ = component of variation due to the additive effects of the genes = $V_{OLO} - \hat{E}$,

$\hat{F}$ = mean covariation of the additive and dominance effects over arrays
= $2V_{OLO} - 4W_{OLOI} - 2(n-2)\hat{E}/n$,

$\hat{H}_1$ = component of variation due to the dominance effects of the genes
= $V_{OLO} - W_{OLOI} + 4V_{ILI} - (3n-2)\hat{E}/n$,

$\hat{H}_2$ = dominance indicated by asymmetry of positive and negative effects of genes
= $4V_{ILI} - 4V_{OLI} - 2\hat{E}$,

$\hat{h}^2$ = square of the dominance effects over all loci in heterozygous phase in all crosses
= $4(M_{LI} - M_{LO})^2 - 4LE$, where
$L = \left(\left(\sum_i \sum_j 1/K_{ij}^2\right) - \left(\sum_i 1/K_{ii}^2\right)\right)\Big/$
$\left(\left(\sum_i \sum_j 1/K_{ij}^2\right)\left(\sum_i 1/K_{ii}^2\right)\right).$

Our components are estimated as:

$\hat{E}$ = expected environmental component of variation = Error A $\times$ Koo, and

Koo = average value of the J_{ij}^2's where $\hat{\sigma}^2 K_{ij}^2$ is the variance of U_{ij}.

The expected values of the components of variation are obtained by least squares analysis and assuming that $\hat{E} = E = E'$. The standard errors of the genetic components are computed using the equation in Hayman (1954).

Example

The following example uses four parents: P_1 ('UI 114'), P_2 (PI 324607), P_3 (PI 271998), and P_4 ('Valley'), to

Table 1. F_1 hybrid and parental data in a four parent diallel cross for the killing temperature, U, and its coefficient of the standard deviation K

		P_1	P_2	P_3	P_4
P_1	U	47.05	46.44	47.40	47.55
	K	1.598	1.959	1.192	1.683
P_2	U	46.44	46.02	48.62	47.00
	K	1.959	1.736	2.056	1.329
P_3	U	47.40	48.62	49.60	46.83
	K	1.192	2.056	2.083	2.250
P_4	U	47.55	47.00	46.83	46.27
	K	1.683	1.329	2.250	1.637

Table 2. Nonlinear analysis of covariance of killing temperature for parents and F_1 hybrids using conductivity test

Source	d. f.	Mean squares
Cross	18	0.2077 *
Plant (cross) − error A	58	0.0774
Error B	897	0.1109
Total	973	

* Significant at 1% level

Table 3. Array covariances, variances, and their sums and differences

Array	W_r	V_r	$W_r + V_r$	$W_r − V_r$
1	0.3510	0.2423	0.5933	0.1087
2	1.3717	1.0773	2.4490	0.2944
3	1.0831	1.2655	2.3486	− 0.1824
4	0.1123	0.3197	0.4320	− 0.2074

Fig. 1. Graph of variance vs covariance (V_r, W_r) of four bean cultivars for heat tolerance

demonstrate the computations necessary to conduct the weighted diallel analysis. Table 1 presents the estimates of U, the killing temperature, which are the data to be analyzed, and the coefficient of its standard deviation, K. The analysis of variance, computed according to the method of Schaff et al. (1987), shows that there are significant differences between the parents and F_1 hybrids for heat tolerance (Table 2). The estimates of the evironmental variance is provided by the Error A or $\hat{\sigma}^2 = 0.0774$ and thus $\hat{E} = 0.0774 \times 3.18$.

The assumptions of the Hayman (1954) and Jinks (1954) diallel model are: (1) diploid segregation, (2) homozygous parents, (3) no differences between reciprocal crosses, (4) independent action of nonallelic genes, (5) no multiple allelism, and (6) genes independently distributed among parents.

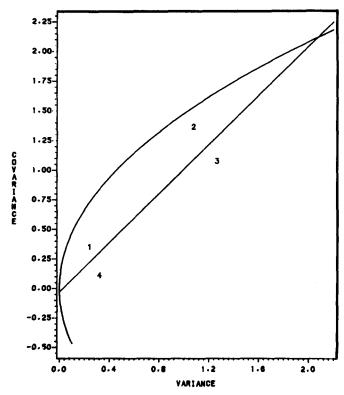The plants used in the diallel had diploid segregation and the parents were homozygous inbreds with no

differences between the reciprocal crosses. Therefore, reciprocals were combined, resulting in ten entries of four parents, and six hybrids. The remaining three assumptions will be tested with the techniques of the Hayman (1954) and Jinks (1954) diallel, weighted for unequal variances of the killing temperature. An examination of the variances and covariances of the arrays reveals that the t-test for heterogeneity of $W_r − V_r$ is nonsignificant, with $t = 0.2185$ (Table 3). Also, the V_r, W_r graph (Fig. 1) has a slope that is not significantly different from unity (slope = 1.0451 $\pm$ 0.3261). Consequently, there is no information to indicate that the assumptions of independent action of nonallelic genes, no multiple allelism, and genes independently distributed between the parents were not valid.

Weighted estimates of the variances and covariances of the parents and crosses were: $V_{OLO} = 2.2039$, $W_{OLOI} = 0.7295$, $V_{ILI} = 0.7262$, $V_{OLI} = 0.2314$, and $M_{LI} − M_{L\hat{0}}^2 = 0.0463$. From these estimates the genetic components were computed (Table 4). Since the diallel analysis had less than 10 entries, these genetic estimates relate only to these parents and crosses and not to a larger population (Hayman 1963). The genetic components indicated that dominance was not significant because $\hat{H}_1$, $\hat{H}_2$, $\hat{F}$, and $\hat{h}^2$ were not significant

Table 4. Genetic variance components for killing temperature of four parent common bean diallel. $\hat{D}$ = additive effects of genes; $\hat{H}_1$ = dominance effects of genes; $\hat{H}_2$ = dominance indicated by asymmetry of positive and negative effects of genes; $\hat{F}$ = covariance of dominance and additive effects; $\hat{h}^2$ = square of the dominance effects over all loci in heterozygous phase in all crosses; $\hat{E}$ = environmental error

Genetic components	Estimate and standard error
$\hat{D}$	1.9579 ± 0.1908 *
$\hat{F}$	1.2439 ± 0.4901
$\hat{H}_1$	1.5757 ± 0.5546
$\hat{H}_2$	1.4874 ± 0.5119
$\hat{h}^2$	0.0040 ± 0.3472
$\hat{E}$	0.2460 ± 0.0853

* Significantly different from zero at 1% level

Table 5. Estimates of genetic parameters for killing temperature in common bean in greenhouse/growth chamber experiments. $(H/D)^{1/2}$ = average degree of dominance; $H_2/4H_1$ = average frequency of positive vs negative alleles; K_D/K_R = $((4D\,H_1)^{1/2}+F)/((4D\,H_1)^{1/2}-F)$ is the ratio of dominant to recessive alleles; h^2/H_2 = the average number of factors involved; Her = $1/4D/(1/4D - 1/4F + 1/4H_1 + E)$

Genetic components	Estimate and standard error
$(H/D)^{1/2}$	0.8971 ± 0.3555
$H_2/4H_1$	0.2360 ± 0.5228
K_D/K_R	2.0964 ± 1.2923
h^2/H_2	0.0027 ± 0.2334
Her	0.5980 ± 0.0866

(Tables 4 and 5), which suggests that most of the genetic variability was due to additive gene action. The coefficient of correlation between the parental degree of dominance $(W_r + V_r)$ and the weighted parental values was nonsignificant $(r = 0.3626 \pm 0.6590)$, indicating

that there are as many positive as negative genes that show dominance. The graph (Fig. 1) shows that parents 1 and 4 had more dominant genes and parents 2 and 3 had more recessive genes.

This analysis can be used in studies where there are unequal numbers of replications per cross, such as that described by Gibori et al. (1978). Although they used an unequal number analysis of variance to obtain estimates of the variance components, they did not use the unequal numbers of replications from the means in the computation of the diallel analysis. Consequently, their analysis is inappropriate whereas the technique described here is more appropriate. A worked example for an unweighted diallel analysis is presented by Aksel and Johnson (1963).

References

Aksel R, Johnson LPV (1963) Analysis of a diallel cross: a worked example. Adv Front Plant Sci 2:37–53

Gibori A, Hillel J, Cahaner A, Ashri A (1978) A 9×9 diallel analysis in peanuts (*A. hypogaea* L.): flowering time, tops' weight, pod yield per plant and pod weight. Theor Appl Genet 53:169–179

Hayman BI (1954) The theory and analysis of diallel crosses. Genetics 39:789–809

Hayman BI (1963) Notes on diallel-cross theory. In: Hanson WD, Robinson HF (eds) Statistical genetics and plant breeding. NAS-NRC 982, pp 571–578

Jinks JL (1954) The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. Genetics 39:767–788

Schaff DA (1984) Screening for, and inheritance of, heat tolerance in common bean (*Phaseolus vulgaris* L.). PhD Dissertation, Kansas State University, Manhattan, Kansas

Schaff DA, Milliken GA, Clayberg CD (1987) A method for comparing nonlinear models for data from a split-plot or repeated measures design. Biomet J (in press)