GENETICAL ANALYSIS OF HALF DIALLELS*

**Anita Dobek[1], Zygmunt Kaczmarek[2], Hanna Kiełczewska[1]
and Tadeusz Łuczkiewicz[3]**

[1]Department of Mathematical and Statistical Methods,
Academy of Agriculture, Poznań, Poland
[2]Institute of Plant Genetics of the Polish Academy of Sciences,
Poznań, Poland
[3]Department of Genetics and Plant Breeding, Academy of Agriculture,
Poznań, Poland

Summary

*Assuming a simple diploid genetical system we show the relations
between various statistics obtained from measurements on the progeny and
the genetical components of Mather. Additionally the analysis of variance
is developed which detects both additive genetic variation and dominance
deviations.*

1. INTRODUCTION

The diallel cross method of investigating the genetical properties of
a group of n homozygous lines has been developed by Hayman (1954a,b) and
Mather and Jinks (1982) and Dobek et al. (1983a,b). They considered a
diallel table consisting of the progeny of n selfed homozygous lines and
their n^2-n crosses. We shall investigate a triangular diallel table
consisting of n selfed homozygous lines and their $n(n-1)/2$ crosses.
From a different point of view, such tables were considered by Jones
(1965) and Pooni et al. (1984).

We define some statistics, which are based on the observations of
$n(n+1)/2$ genotypes, and show their relations with Mather's genetical
components. These components are useful for detecting the additive and

dominance variation and for describing the genetical properties of parents. Additionally, analysis of variance and tests of some interesting hypotheses, concerning additive and dominance effects, are presented.

The estimators of Mather's parameters given in this paper are different from those obtained by Ferreira (1985). The reason is that Ferreira used statistics analogous with statistics for complete diallels. Our statistics are designed specially for half diallel situation.

2. GENETICAL MODEL

To find the relations between the appropriate statistics and the genetical components it is necessary to assume the absence of: maternal effects, epistasis, linkage and multiple allelism.

Suppose that the investigated character is controlled by r genes. According to the Mather's notation the effects of three possible genotypes $A_k A_k$, $A_k a_k$, $a_k a_k$ in the k-th locus are d_k, h_k, $-d_k$, where $d_k > 0$. Let Θ_k be a variable which takes for the three genotypes the values 1, 0, -1 respectively. We will assume that in the diallel cross the frequency of parents with genotype $A_k A_k$ is u_k and with genotype $a_k a_k$ is v_k, where $u_k + v_k = 1$.

Let us represent the genotype of the i-th parental line by $\Theta_i = \{\Theta_{i1}, \Theta_{i2}, \ldots, \Theta_{ir}\}$ where Θ_{ik} takes the values according to the genotype of the i-th line in the k-th locus, $i=1,2,\ldots,n$, $k=1,2,\ldots,r$. Then, as it is shown by Hayman (1954a), the expected value of genotype obtained as a result of cross $\Theta_i \times \Theta_j$ is

$$\alpha_{ij} = \frac{1}{2} \sum_{k=1}^{r} [d_k(\Theta_{ik} + \Theta_{jk}) + h_k(1 - \Theta_{ik}\Theta_{jk})]$$

and the genotypic value of the i-th parental line is $\alpha_{ii} = \sum_{k=1}^{r} d_k \Theta_{ik}$.

3. STATISTICS OF THE TRIANGULAR DIALLEL TABLE

The means of the observations from an experiment with the mating combinations of a diallel cross can be arranged in a triangular table of the form

$$
\begin{array}{cccc}
y_{11} & y_{12} & \cdots & y_{1n} \\
 & y_{22} & \cdots & y_{2n} \\
 & & & y_{nn}
\end{array}
$$
,

where y_{ij}, $i \le j = 1,2,\ldots,n$, are normally, independently distributed with the expected value α_{ij} and variance σ^2. We assume that the experiment is laid out in a complete block design with b blocks.

The following statistics will be useful in estimating the genetical parameters

- parental mean

$$\bar{y}_P = \frac{1}{n}y_P \quad , \quad \text{where} \quad y_P = \sum_{i=1}^{n} y_{ii} \quad ,$$

- mean of the diallel table

$$\bar{y}_{..} = y_{..}/[n(n+1)/2] \quad , \quad \text{where} \quad y_{..} = \sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij} \quad ,$$

- variance of the parents

$$V_P = \frac{1}{n-1}\sum_{i=1}^{n} (y_{ii} - \bar{y}_P)^2 \quad ,$$

- covariance between the parents and the means of their families

$$W_{PF} = \frac{1}{n-1}\sum_{i=1}^{n} (y_{ii} - \bar{y}_P)[(y_{i.} + y_{ii})/(n+1) - \bar{y}_{..}] \quad ,$$

- average variance of the families

$$V_{MF} = \frac{1}{n(n-1)}\left\{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} [(y_{ij}-\bar{y}_{i.})^2 + (y_{ij}-\bar{y}_{j.})^2] + \sum_{i=1}^{n} (y_{ii}-\bar{y}_{i.})^2\right\} \quad ,$$

- variance of the means of the families

$$V_F = \frac{1}{n-1}\sum_{i=1}^{n} [(y_{i.} + y_{ii})/(n+1) - \bar{y}_{..}]^2 \quad ,$$

where $y_{i.} = \sum_{j=1}^{n} y_{ij}$ with $y_{ij} = y_{ji}$.

The expected values of the statistics given above may be written in a form similar to that used by Mather and Jinks (1982) as

$$E(V_P) = \frac{1}{n-1}D + \sigma^2$$

$$E(V_F) = \frac{n}{4(n-1)(n+1)}[(n+2)^2D + n^2H_1 - n^2H_2 - n(n+2)F] + \frac{n+2}{(n+1)^2}\sigma^2 \quad ,$$

$$E(V_{MF}) = \frac{n}{4(n-1)}(D + H_1 - F) + \sigma^2 \quad ,$$

$$E(W_{PF}) = \frac{n}{2(n^2-1)}[(n+2)D - \frac{1}{2}n\ F] + \frac{2}{n+1}\sigma^2 \quad ,$$

$$E(\bar{y}_{..} - \bar{y}_P) = \frac{n}{2(n+1)}h \quad ,$$

where

$$D = \sum_{k=1}^{r} d_k^2(1-w_k^2) \quad ,$$

$$F = 2\sum_{k=1}^{r} d_k h_k w_k(1-w_k^2) \quad ,$$

$$H_1 = \sum_{k=1}^{r} h_k^2(1-w_k^2) \quad ,$$

$$H_2 = \sum_{k=1}^{r} h_k^2(1-w_k^2)^2 \quad ,$$

$$h = \sum_{k=1}^{r} h_k(1-w_k^2) \quad ,$$

and $w_k = u_k - v_k$.

By equating the statistics to their expected values and taking as the estimator of σ^2 the mean square error in the randomized block analysis MS$_e'$ divided by b, i.e. $\hat{\sigma}^2 = MS_e'/b = MS_e$, we obtain

$$\hat{D} = \frac{n-1}{n}(V_P - \hat{\sigma}^2) \quad ,$$

$$\hat{F} = \frac{2(n-1)}{n^2}[(n+2)V_P - 2(n+1)W_{PF} - (n-2)\hat{\sigma}^2] \quad ,$$

$$\hat{H}_1 = \frac{n-1}{n^2}[4n\,V_{MF} + (n+4)V_P - 4(n+1)W_{PF} - (5n-4)\hat{\sigma}^2] \quad ,$$

$$\hat{H}_2 = \frac{4(n-1)}{n^3}[n^2 V_{MF} - (n+1)^2 V_F - V_P + 2(n+1)W_{PF} - (n^2-n+1)\hat{\sigma}^2] \quad ,$$

$$\hat{h}^2 = \frac{4(n+1)^2}{n^4}(\bar{y}.. - \bar{y}_P)^2$$

4. ANALYSIS OF VARIANCE AND TESTING THE HYPOTHESES

Similarly to the analysis of variance of diallel table given by Hayman (1954b) we can express the expected values of mean squares for general and specific combining abilities in terms of genetical parameters. It is possible to divide the sum of squares for specific combining ability into the component s_1 useful for testing and the residual s_2.

By defining the sums of squares for general combining ability (g.c.a.), for specific combining ability (s.c.a.) and for the components s_1 and s_2 as

$$SS_g = \frac{1}{(n+2)}[\sum_{i=1}^{n}(y_i. + y_{ii})^2 - n(n+1)^2\bar{y}..^2] \quad ,$$

$$SS_s = \sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij}^2 - \frac{1}{n+2}\sum_{i=1}^{n}(y_i + y_{ii})^2 + \frac{2(n+1)}{2(n+2)}\bar{y}..^2 \quad ,$$

$$SS_{s_1} = \frac{n(n+1)}{n-1}(\bar{y}.. - \bar{y}_P)^2 \quad ,$$

$$SS_{s_2} = SS_s - SS_{s_1} \quad ,$$

we obtain the analysis of variance table shown in Table 1. In this table

$$\nu_g = n-1 \ , \quad \nu_s = \frac{n(n-1)}{2} \ , \quad \nu_e = \frac{(b-1)(n-1)(n+2)}{2} \quad \text{and} \quad \nu_{s_2} = \frac{(n+1)(n-2)}{2} \ ,$$

and the mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom.

Table 1. Analysis of variance

Source of variation	Degrees of freedom	Mean squares	Expected mean squares
g.c.a.	ν_g	MS_g	$\frac{1}{4(n+2)}[(n+2)^2 D + n^2 H_1 - n^2 H_2 - n(n+2)F] + \sigma^2$
s.c.a.	ν_s	MS_s	$\frac{1}{n+2}[H_1 + \frac{n^2}{4(n+1)}H_2] + \sigma^2$
component s_1	1	MS_{s_1}	$\frac{n^2}{4(n+1)}h^2 + \sigma^2$
residual s_2	ν_{s_2}	MS_{s_2}	–
error	ν_e	MS_e	σ^2

It is convenient to begin the testing of hypotheses by the following hypothesis:

$$H_o : H_1 + \frac{n^2}{4(n+1)}H_2 = 0 \quad , \tag{1}$$

which is true when all $h_k = 0$ what clearly detects mean square dominance. The hypothesis (1) is rejected when the ratio $MS_s/MS_e \geq F_{\alpha;\nu_s,\nu_e}$. In the case of rejection of the hypothesis (1) it is reasonable to test the hypothesis

$$H_o : h^2 = 0 \tag{2}$$

by comparing MS_{s_1}/MS_e to $F_{\alpha;1,\nu_e}$. If (2) is not true there exists the majority of positive or negative dominants. If the hypothesis (1) is true it is possible to detect the existence of additive effects by testing the hypothesis

$$H_o : (n+2)^2 D + n^2 H_1 - n^2 H_2 - n(n+2)F = 0 \quad , \tag{3}$$

which is equivalent to the hypothesis $D=0$. This hypothesis is rejected when the ratio $MS_g/MS_e \geq F_{\alpha;\nu_g,\nu_e}$.

5. INTERPRETATION OF GENETICAL COMPONENTS

The genetical components of Mather have the following interpretation:

D = component of variation due to the additive effects of the genes,

F = mean covariation of the additive and dominance effects over arrays,

H_1 = component of variation due to the dominance effects of the genes,

H_2 = dominance indicated by asymmetry of positive and negative effects of genes,

h^2 = square of dominance effects over all loci in heterozygous phase in all crosses.

The estimated genetical components can be used for calculating the following:

- the number of groups of genes which control the character and exhibit dominance, as $\hat{h}^2/\hat{H}_2$,
- the average degree of dominance, as $\sqrt{\hat{H}_1/\hat{D}}$,
- the frequency product $(u_k v_k)$ at loci exhibiting dominance, as $\hat{H}_2/4\hat{H}_1$,
- the ratio of dominant to recessive alleles in the parents, as $(\sqrt{4\hat{D}\hat{H}_1} + \hat{F})/(\sqrt{4\hat{D}\hat{H}_1} - \hat{F})$.

Other parameters which can be calculated on the basis of the estimated genetical components are narrow and broad heritabilities, namely

$$h_n^2 = \frac{\hat{D} + \hat{H}_1 - \hat{H}_2 - \hat{F}}{\hat{D} + \hat{H}_1 - \frac{1}{2}\hat{H}_2 - \hat{F} + 2\hat{\sigma}^2} ,$$

$$h_b^2 = \frac{\hat{D} + \hat{H}_1 - \frac{1}{2}\hat{H}_2 - \hat{F}}{\hat{D} + \hat{H}_1 - \frac{1}{2}\hat{H}_2 - \hat{F} + 2\hat{\sigma}^2} .$$

6. EXAMPLE

To illustrate the theory presented in this paper let us consider the statistical - genetical analysis of an experiment with $n=5$ inbred lines and $n(n-1)/2 = 10$ F_1 hybrids of sunflower obtained from a diallel cross. The experiment, performed at the Academy of Agriculture in Poznań in 1980, was laid down in randomized block design with $b=4$ replications. The observed trait was the length of internodes.

Table 2 contains the average values of the length of internodes for the 15 analysed genotypes.

From the experiment we obtain $\hat{\sigma}^2 = 0.203$, $\bar{y}_p = 7.86$ and $\bar{y}_{.\lambda} = 8.63$. The estimates of the genetical Mather's parameters are: $\hat{D} = 4.29$, $\hat{F} = 3.73$, $\hat{H}_1 = 4.66$, $\hat{H}_2 = 3.82$ and $\hat{h}^2 = 3.44$. We can also provide the analysis of variance given in Table 3.

Table 2. Mean values of the length of internodes (in cm.) for parental lines and F_1 hybrids of sunflower

y_{ij}		Male					Sum y_i.
		1	2	3	4	5	
	1	11.8	7.4	9.7	9.6	10.4	48.9
	2		6.5	9.3	9.8	8.1	41.1
Female	3			6.6	9.2	7.5	42.3
	4				8.3	9.2	46.1
	5					6.1	41.3

Table 3. Analysis of variance for the length of internodes of 5 parental lines and 10 F_1 hybrids of sunflower

Source of variation	Degrees of freedom	Mean squares	F	$F_{o.o5}$	$F_{o.o1}$
g.c.a.	4	4.692	23.08	2.59	3.80
s.c.a.	10	1.720	8.46	2.06	2.77
components s_1	1	4.485	22.07	4.07	7.27
residual s_2	9	1.413	-	-	-
experimental error	42	0.203			

As it is easily seen all the hypotheses tested in the analysis of variance are rejected at $\alpha = 0.01$. It means that there exists a dominance (rejection of hypothesis 1) and it is a one way dominance (rejection of hypothesis 2).

Based on the estimates of Mather's parameter we have obtained the degree of dominance equal to 1.04, number of groups of genes which control the character and exhibit dominance equal to 1 (exactly 0.90), frequency product (uv) at loci exhibiting dominance equal to 0.20, the ratio of dominant to recessive alleles equal to 2.43 and coefficients of narrow and broad heritability equal to 0.38 and 0.89, respectively.

REFERENCES

Dobek, A., Kaczmarek, Z., Kiełczewska, H., Łuczkiewicz, T. (1983a). Podstawy teoretyczne analizy krzyżówek diallelicznych. Część I. Analiza wariancji pełnej tablicy diallelicznej wyników doświadczenia o blokach kompletnych. *Biuletyn Instytutu Hodowli i Aklimatyzacji Roslin* **151**, 9-18.

Dobek, A., Kaczmarek, Z., Kiełczewska, H., Łuczkiewicz, T. (1983b). Podstawy teoretyczne analizy krzyżówek diallelicznych. Część II. Analiza genetyczna pełnej tablicy diallelicznej wyników doświadczenia o blokach kompletnych. *Biuletyn Instytutu Hodowli i Aklimatyzacji Roślin* 151, 19-29.

Ferreira, P.E. (1985). On Jinks-Hayman analysis of half diallels. *Rev. Brasil. Genet.* VII, 1, 149-155.

Hayman, B.I. (1954a). The theory and analysis of diallel crosses. *Genetics* 39, 789-809.

Hayman, B.I., (1954b). The analysis of variance of diallel tables. *Biometrics* 10, 235-244.

Jones, R.M. (1965). Analysis of variance of half diallel table. *Heredity* 20, 117-121.

Mather, K., Jinks, J.L. (1982). *Biometrical Genetics.* 3rd edn, Chapman and Hall, London.

Pooni, H.S., Jinks, J.L., Singh, J.L. (1984). Methods of analysis and estimation of the genetic parameters from a diallel set of crosses. *Heredity* 52, 243-253.

ANALIZA GENETYCZNA LINII RODZICIELSKICH I MIESZAŃCÓW
OTRZYMANYCH W WYNIKU PROSTEGO KRZYŻOWANIA DIALLELICZNEGO

Streszczenie

Przy założeniu diploidalnego systemu dziedziczenia, przedstawiono związki pomiędzy pewnymi statystykami uzyskanymi na podstawie obserwacji potomstwa otrzymanego w wyniku prostego krzyżowania diallelicznego a komponentami genetycznymi Mathéra. Podana ponadto analiza wariancji pozwala testować hipotezy dotyczące addytywnego i dominacyjnego działania genów.