ANALYSIS OF A 6×6 DIALLEL CROSS FOR EGG PRODUCTION IN *DROSOPHILA MELANOGASTER*

By H. P. SINGH, S. S. PRABHU, PREM NARAIN and SWASTIKA NEGI
Indian Veterinary Research Institute, Izatnagar, U.P.

Diallel crosses have been employed by Jinks (1954, 1955, 1956) in the study of certain quantitative characters in plants. Patricia Cooke and Mather (1962) utilised them for estimating the components of continuous variation in sternopleural bristles in *Drosophila melanogaster*. In the present study are reported the results of a 6×6 diallel cross for egg production carried out in *Drosophila melanogaster*. As far as is known to the authors, this is the first time such a study has been attempted with a trait like egg production in this fly.

Material and Method

Full details of the lines and their average egg production prior to starting the experiment are presented in Table 1.

Table 1. *Details of the experimental lines used in the diallel cross.*

Line No.	Status	Original parent	Mean egg production
A	27H-1	P-4 (N-18)	65·5
B	27H-3	P-5 (N-18)	57·4
C	25H-2A	24 (N-6)	71·7
D	25L-1	21 (N-6)	33·2
E	27L-2	P-4 (N-18)	49·5
F	26L-4A	P-1 (N-18)	28·6

In Table 1, 27H-1 means high line 1, sib-mated and selected for 27 generations for high egg production. 25L-1 means low line 1 sib-mated and selected for 25 generations for low egg production and so on. The symbols P-4, 24, etc. in the 3rd column stand for the number of the fly, the stock number being given in parenthesis. N stands for *Naibasti*, a place close to the Institute.

Prior to starting the experiment, the lines had been relaxed and allowed to multiply in a culture bottle for about 9 months, so that the populations in the bottles at the beginning of the diallel matings were in equilibrium conditions. Though there were 31 lines available for the cross, because of the heavy labour involved in carrying through the work, only six lines were ultimately selected; three of these *viz*., A, B and C, had higher egg production than the other three, D, E and F lines. The former set came from lines selected for high egg production and the latter from lines selected for low egg production.

Crosses were made in the following way:

Sets of 4 virgins were made from each of the 6 selected inbred lines. From a given line, each set of virgins was mated to a single male belonging to each of the 6 inbred lines. There were four such replications, and with reciprocal matings included they gave a total of $6\times4\times6\times4=576$ matings in all.

The technique adopted for collecting eggs and counting them was the same as described by Prabhu (1960). The first generation provided the data concerning the performance of the six lines, as also those of the F_1 between them.

The experimental material was kept throughout in a constant temperature cabinet maintained at a temperature of 77°F with a sensitivity of ±1°F.

Analytical Procedure

The procedure followed in the statistical treatment of the data is that suggested by Hayman (1954).

Firstly, variances within families of parents and F_1s were tested for genotype-environment interaction. In the present case, a trend existed between all the family means and variances; hence, the raw data was transformed to logarithms.

Secondly, from the diallel tables of reciprocal means in terms of the transformed values, V_r and W_r ($r=1, 2\ldots6$), defined below, were computed for each replication.

V_r — the variance of the rth array,

W_r — the covariance between the parents and their offspring in the rth array.

$(W_r—V_r)$ were then tested for heterogeneity performing the analysis of variance of $(W_r—V_r)$. The two values of $(W_r—V_r)$ for $r=3$ and 4 deviated markedly from the common values of others, which as per the analysis, required their deletion. In this way, the values of progenies of crosses C×D and D×C had to be omitted. The reason why these values differed from others perhaps lay in the fact that C and D lines came from N-6 stock, while all the others were from N-18 stock. Using the missing plot technique, the values were found and introduced to secure uniformity in $W_r—V_r$ values.

Thirdly, the uniform $(W_r—V_r)$ diallel table was subjected to the analysis of variance to test the significance of the genetical parameters of variation, and for providing the estimate of error (E). The total variance was split up as follows:

(a) = variation between the mean effects of each parental line,

(b) = variation in the reciprocal sums not ascribable to (a),

(c) = average maternal effects of each parental lines,

(d) = variation in the reciprocal differences not ascribable to (c),

R = variation due to replication and the interaction of the above main effects with replication.

Further, the sum of squares (b) was split up into the following components:

(b_1) = mean dominance deviation,

(b_2) = further dominance deviation due to the rth parent,

(b_3) = remaining discrepancy in the rsth reciprocal sum,

Fourthly, from each of the four 6×6 diallel tables, the following statistics were calculated:

$(m_{L1}-m_{L0})^2$ = the square of the difference between the mean of the parents and the mean of their 36 progenies.

V_{0L0} = the variance of the parents,

V_{0L1} = the variance of the means of the arrays,

V_{1L1} = the mean of the variances of arrays, and

W_{0L01} = covariance between the parents and the means of their offspring.

These statistics, together with an estimate of E, provided the least square solution of the various genetic components $\hat{D}$, $\hat{F}$, $\hat{H}_1$, $\hat{H}_2$, $\hat{h}^2$ and $\hat{F}_r$.

$$\hat{D} = V_{0L0} - \hat{E}$$

$$\hat{F} = 2V_{0L0} - 4W_{0L01} - \frac{2\times4}{6}\hat{E}$$

$$\hat{H}_1 = V_{0L0} - 4W_{0L01} + 4V_{1L1} - \frac{18-2}{6}\hat{E}$$

$$\hat{H}_2 = 4V_{1L1} - 4V_{0L1} - 2\hat{E}$$

$$h^2 = 4(m_{L1} - m_{L0})^2 - \frac{4\times5}{36}\hat{E}$$

$$\hat{F}_r = 2(V_{0L0} - W_{0L01} + V_{1L1} - W_r - V_r) - \frac{2\times4}{6}\hat{E}.$$

$(r=1, 2....6)$

Fifthly, the following ratios were further calculated by using the estimates $\hat{D}$, $\hat{F}$, $\hat{H}_1$, $\hat{H}_2$, and h^2.

$$\frac{\hat{H}_1}{\hat{D}}, \quad \frac{\hat{H}_2}{4\hat{H}_1}, \quad \frac{(4D\hat{H}_1)^{\frac{1}{2}}+\hat{F}}{(4\hat{D}\hat{H}_1)^{\frac{1}{2}}-\hat{F}} \text{ and } \frac{h^2}{\hat{H}_2}.$$

Sixthly, the residual mean square (S^2) for calculation of standard errors of the components was obtained from the sum of squares of deviations of observed, from expected values of V_{0L0}, V_{0L1}, V_r and W_r $(r=1....6)$ and divided by $57-12=45$, the degrees of freedom for error. This was based on the consideration that the four diallel tables would provide 57 statistics, and 12 constants are fitted to the data. The standard errors were calculated from the following formulae:

$$\text{Var. } D = \frac{S^2}{4n^5}(n^5+n^4)$$

$$\text{Var. } \hat{F} = \frac{S^2}{4n^5}(4n^5+20n^4-16n^3+16n^2)$$

$$\text{Var. } \hat{H}_1 = \frac{S^2}{4n^5}(n^5+41n^4-12n^3+4n^2)$$

$$\text{Var. } \hat{H}_2 = \frac{S^2}{4n^5}.36n^4$$

$$\text{Var. } h^2 = \frac{S^2}{4n^5}(16n^4+16n^2-32n+16)$$

$$\text{Var. } \hat{E} = \frac{S^2}{4n^5}.n^4$$

where $n=6$.

RESULTS

In Table 2 are presented the mean egg production of the six parents along with their different F_1 progenies. The figures are in original units.

Table 2. *Mean egg production*

		FEMALES					
		A	B	C	D	E	F
MALES	A	78·8	94·7	90·2	91·9	90·0	87·7
	B	92·6	84·0	89·2	80·6	90·2	83·5
	C	79·6	86·9	88·2	58·2	94·1	93·4
	D	93·9	91·2	67·9	92·9	96·1	97·3
	E	85·6	87·4	107·0	92·6	94·2	97·4
	F	95·2	86·8	93·2	87·7	89·4	55·0

It is instructive to note the mean egg productions now observed in the parental stock and compare them with the figures found on the day the six lines were relaxed and mass cultured. It shows that as a result of relaxing and mass culturing over a period of 9 months, the lines appeared to have regained their fecundity, particularly the low lines D, E and F, so that with the exception of F, they appear to be of the same level now. Further an examination of the parental means and F_1 means, also shows that in case of A, B, and F lines, the F_1 means, without exception were higher than P means. In C and E, with one exception, they were also higher; while in D, without any exception, the P value was higher than the F_1 mean values. A test of significance carried out on overall P and F_1 means gave the following results:

Parental mean	82·2
F_1 mean	88·0
$(F_1—P)$ mean	5·8
Significance	$<·05$

On an average, heterosis was present in F_1s.

The variances within families of parents and F_1s calculated from the four replicated values of each entry in the 6×6 diallel table are shown in Table 3.

Table 3. *Variances within families*

		FEMALES					
		A	B	C	D	E	F
MALES	A	637·9	279·7	65·2	245·2	343·5	492·3
	B	350·6	149·0	295·0	203·3	148·3	125·8
	C	286·0	286·5	326·5	438·3	525·1	242·3
	D	357·3	635·0	218·1	272·7	290·4	119·6
	E	260·2	512·5	150·2	376·6	244·6	376·2
	F	302·1	637·8	597·5	309·1	348·3	275·0

There was heterogeneity present in the variances of the parents and their F_1s, indicating the existence of genotype-environment interaction. Further, as a trend existed between the means and the variances in the F_1 table, the original data was transformed to logarithms before further analysis.

Table 4 contains the mean of $\log_e x$ of progenies (females) for each of the four replicates, where x denotes egg production per day calculated on the basis of 3-day egg production. In the Table are also presented the variances V_r and the covariances W_r, calculated from the diallel table of mean reciprocals.

Table 4. *Mean of* $\log_e$x

FEMALES

		A	B	C	D	E	F	W_r	V_r
					Replication—1				
MALES	A	4·41	4·52	4·87	4·51	4·53	4·39	·00088	·00594
	B	4·60	4·46	4·48	4·36	4·48	4·46	·00180	·00558
	C	4·38	4·43	4·32	4·01	4·68	4·68	—·01804	·05376
	D	4·60	4·38	3·86	4·41	4·52	4·65	—·00226	·05529
	E	4·39	4·58	4·65	4·45	4·31	4·58	—·00148	·01430
	F	4·59	4·31	4·53	4·50	4·55	4·19	·00726	·02446
					Replication—2				
MALES	A	4·16	4·60	4·41	4·45	4·25	4·48	—·00046	·02170
	B	4·37	4·23	4·43	4·38	4·30	4·45	—·00834	·00938
	C	4·19	4·39	4·22	4·11	4·28	4·58	—·01350	·01638
	D	4·57	4·51	4·40	4·59	4·48	4·47	·00994	·01206
	E	4·42	4·29	4·59	4·44	4·56	4·55	·01176	·00854
	F	4·61	4·44	4·55	4·38	4·32	3·92	·04524	·05562
					Replication—3				
MALES	A	4·31	4·54	4·53	4·46	4·57	4·39	·00580	·00790
	B	4·58	4·45	4·52	4·41	4·60	4·48	—·00436	·00262
	C	4·51	4·42	4·57	3·99	4·47	4·31	·00306	·02710
	D	4·56	4·43	4·28	4·44	4·60	4·57	—·00628	·02148
	E	4·46	4·42	4·72	4·51	4·65	4·56	·00162	·00290
	F	4·60	4·56	4·65	4·33	4·66	3·99	·04660	·04834
					Replication—4				
MALES	A	4·25	4·46	4·46	4·59	4·53	4·47	·00970	·01266
	B	4·45	4·50	4·46	4·35	4·57	4·23	·01646	·00380
	C	4·31	4·49	4·53	3·86	4·57	4·48	·01886	·04514
	D	4·58	4·59	4·17	4·40	4·59	4·58	·02724	·04906
	E	4·46	4·45	4·67	4·64	4·45	4·54	·00192	·00486
	F	4·32	4·46	4·12	4·53	4·54	3·79	·06460	·07852

Analysis of variance of $(W_r - V_r)$ carried out to test the heterogeneity of $(W_r - V_r)$ gave the following results.

	d.f.	M.S.
Strains	5	·00115*
Replicates	3	·00038
Residual	15	·00032

An examination of the $(W_r - V_r)$ values in the four replicates showed that the crosses involving C and D strains gave on an average higher values, which probably accounted for the heterogeneity observed. The actual values are presented in Table 5 which

Table 5

Replication / Line	1	2	3	4	Total
A	—·00506	—·02216	—·00210	—·00296	—·03228
B	—·00378	—·01772	—·00698	·01268	—·01580
C	—·07180	—·02988	—·02404	—·02628	—·15200
D	—·05755	—·00028	—·02776	—·07630	—·16133
E	—·01578	—·00322	—·00128	—·00294	—·01678
F	—·01720	—·01038	—·00174	—·01392	—·04324
Total	—·17117	—·07664	—·06390	—·10972	—·42143

will clarify the position. In an attempt to restore homogeneity, as suggested by Hayman (1954), these crosses were excluded from each of the 4 replications and estimates of their values obtained using the missing plot technique. In this way, the replication-interaction sum of squares was minimised, but not *totally* removed. This was verified by employing the scaling test as suggested by Jinks (1955). In this way, care was taken to see that all the significant non-allelic interaction was not omitted, while omitting the values of progenies corresponding to the crosses C × D and D × C in the diallel table. The estimated values were as under:

Replication	C × D	D × C
1	4·52	4·66
2	4·59	4·34
3	4·40	4·53
4	4·62	4·46

The new entries of $(W_r - V_r)$ in Table 5 for the lines C and D calculated on the basis of the above estimated values come to:

	R_1	R_2	R_3	R_4	Total
C	—·01866	—·01576	·00452	—·01088	—·04078
D	—·01486	·00614	·00164	—·01178	—·01886

The analysis of variance of $(W_r - V_r)$ after introduction of the new values is summarised in Table 6.

Table 6. *Analysis of Variance*

Source of variation	d.f.	M.S.
Strains	5	·00004
Replication	3	·00016
Residual	15	·00013
Total	23	

No heterogeneity is now present, establishing the uniformity of $(W_r - V_r)$ values. Figure 1 shows the points (W_r, V_r), the limiting parabola $W_r^2 = V_{0L0} V_r$, and the line of unit slope through the adjusted mean point (V_{1L1}, W_{0L01}). The unadjusted points lie well off this line.

The four diallel tables containing the estimated values of the progenies of crosses C×D and D×C were subjected to the analysis of variance in the way described by Hayman (1954). In Table 7 are summarised the results showing the significance of the various components of variation worked out.

Table 7. *Components of Variance*

Sources of variation	D.F.	M.S.	F ratio
(a) $D-F+H_1-H_2$	5	0·07346	5·97**
(b) H_2	15	0·07210	5·86**
(c) Reciprocal	5	0·02294	1·86
(d) Differences	10	0·01477	1·20
Total (t)	35		
R (Replications)	3	0·06160	5·00**
R×(a)	15	0·01962	
R×(b)	45	0·01060	
R×(c)	15	0·01049	
R×(d)	30	0·01211	
R×(t)	105	0·01230	

Bartlett's test of heterogeneity of the four interaction variances gives $\chi_3^2 = 2{\cdot}65$, so that in this case all the interactions may be pooled to give (Rt) as a common variance. Comparison with this provides the F value in the last column of Table 7.

The significance of (a) indicates significant additive genetic effects.

The significance of (b) indicates significant dominant effects.

The non-significance of (c) and (d) indicates the absence of significant maternal effect on egg production.

The component (b) can be further split up as below (Table 8):

Table 8. *Splitting up of "b" component*

Source of variation	d.f.	M.S.	F ratio
(b_1)	1	0·46510	139·66**
(b_2)	5	0·09374	4·63**
(b_3)	9	0·01640	2·72*
R × (b_1)	3	0·00333	
R × (b_2)	15	0·02025	
R × (b_3)	27	0·00603	

The significance of (b_1) shows that the dominant effects present were largely *unidirectional*. The progeny mean is greater than the parental mean so that the dominance is in the direction of more egg production.

The significance of (b_2) indicates an asymmetry in the distribution of allels in the parents. The significance of (b_3) also indicates dominance.

The estimate of E was obtained by pooling over all the individual interaction mean squares in the replications.

The values of statistics enumerated in item four of the analytical procedure were also obtained for each of the four diallel tables and are given in Table 9.

Table 9

Replication	V_{0L0}	W_{0L01}	V_{1L1}	V_{0L1}	$(m_{L1}-m_{L0})^2$
R_1	·00948	·00008	·01263	·00095	·0225
R_2	·06492	·00951	·01895	·00140	·0169
R_3	·05435	·01025	·01124	·00251	·0081
R_4	·07712	·01501	·09980	·00628	·0196
Average	·05146	·00871	·01570	·00278	·0168

The mean values of these statistics together with the mean values of W_r and V_r (r=1, 2....6) and the estimate of E were used to obtain the values of the genetic components D, F, H_1, H_2 and h^2. The standard errors of these were also calculated as per the formulae given under item six of the analytical procedure. The estimates along with their standard errors are given in Table 10.

Table 10

Statistics	$\hat{D}$	$\hat{F}$	$\hat{H}_1$	$\hat{H}_2$	$\hat{h}^2$	$\hat{E}$
Value	·03916	·05168	·04662	·02708	·06029	·01230
Standard error	·00707	·01673	·01732	·01549	·01049	·00316
"t"	5·53**	3·09**	2·69**	1·74	5·74**	3·89**

$\hat{H}_1$ is significant, but $\hat{H}_2$ is not so in this test, though the more reliable test provided by the analysis of variance indicated the presence of dominance.

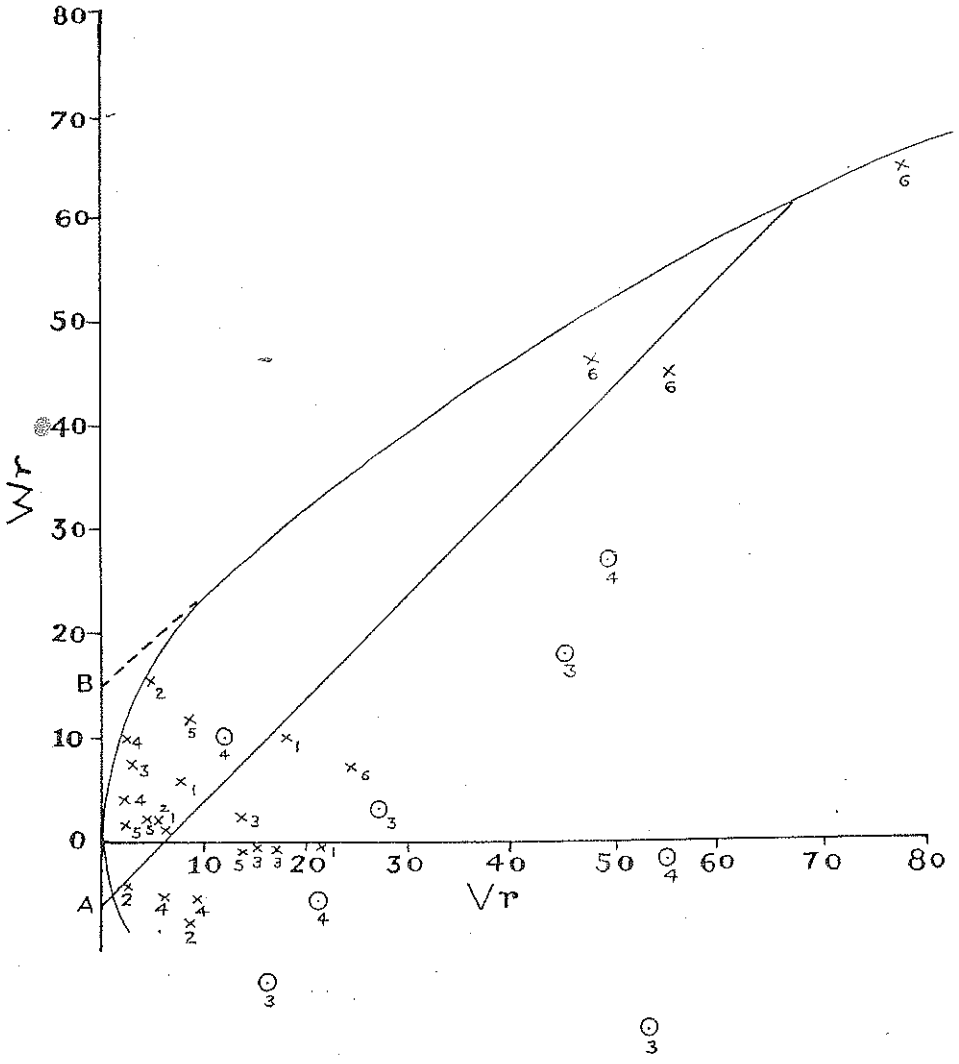

Fig. 1. (W_r, V_r) graph for egg production in replicated crosses of six lines of *Drosophila melanogaster*.

The value of $\left[\frac{\hat{H}_1}{\hat{D}}\right]^{\frac{1}{2}} = 1{\cdot}091$, gives an estimate of the mean degree of dominance over all loci. $\frac{\hat{H}_1}{\hat{D}}$ also estimates $\frac{AB}{OB}$ in the (V_r, W_r) graph, where A is the point where the straight line of unit slope cuts the axis, O is the origin and B is the point where the tangent to the limiting parabola $W_r^2 = {\cdot}051464\ V_r$ cuts the W_r axis. After adjusting the measurements of the progeny of the crosses of lines C and D, the points lie nearer the straight line of unit slope inside the limiting parabola. The unadjusted points (circled) lie well off this line.

The value of $\frac{\hat{H}_2}{4\hat{H}_1} = 0{\cdot}145$ shows that H_1 is significantly different from H_2, which is supported by the significance of (b_2) in the analysis of variance, referred to above. This, together with the observation that $H_1 > H_2$, indicates that the positive and negative alleles at loci exhibiting dominance are not in equal proportions in the parents.

The value of $\frac{(4D\hat{H}_1)^{\frac{1}{2}} + \hat{F}}{(4\hat{D}\hat{H}_1)^{\frac{1}{2}} - \hat{F}}$ is found to be 4·061. This indicates that there are more dominant genes than recessive ones. For every one recessive gene affecting egg production, we can expect four dominant genes.

The value $\frac{\hat{h}^2}{\hat{H}_2}$ has been obtained as 2·226 which shows that there are at least two genes out of all the genes controlling egg production, which exhibit dominance.

The order of dominance of the parents determined by the value of $(W_r + V_r)$ is DBECAF, and the order of egg production is EDCBAF, parent F having the lowest egg production and carrying the least dominants.

CONCLUSIONS

1. There is a considerable amount of interaction between the environment and the genetic factors controlling egg production in *Drosophila melanogaster*.

2. The genes affecting egg production exhibit both additive genetic effects and dominant deviations. They also show non-allelic interactions. The non-allelic interaction between C and D results in low egg production.

3. An asymmetry in the distribution of positive and negative alleles exists in the parents, which was understandable as three of the six lines included in the trial were selected for low egg production and the other three for high egg production.

4. There are probably present more dominant than recessive genes affecting egg production in the parental stock used in these studies; the ratio of recessives to dominants being of the order of 1:4. Low egg production is due to recessive genes, high egg production to their dominant allels.

5. Dominance deviations present were largely unidirectional.

6. The order of dominance of the six lines used was D, B, E, C, A, F.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. J. L. Jinks, University of Birmingham, for having kindly gone through the draft of the paper and for valuable suggestions. They are indebted to the Director, Indian Veterinary Research Institute, Izatnagar, U.P. for providing all facilities.

REFERENCES

COOKE, PATRICIA AND MATHER, K. (1962). Estimating the components of continuous variation, II. Genetical. *Heredity*, **17**, 211–236.

HAYMAN, B. I. (1954). 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.

JINKS, J. L. (1955). A survey of the genetical basis of heterosis in a variety of diallel crosses. *Heredity*, **9**, 233–238.

JINKS, J. L. (1956). The F_2 and back cross generations from a set of diallel crosses. *Heredity*, **10**, 1–30.

PRABHU, S. S. (1960). Progress report of the scheme for "Fundamental study of quantitative inheritance" at the IVRI-Izatnagar, U.P. (unpublished).