AN EFFECT OF GENE ARRANGEMENT ON THE RECOMBINATION FRACTION IN *DROSOPHILA MELANOGASTER*

P. A. PARSONS
Department of Genetics, Cambridge

Received 26.iv.56

1. INTRODUCTION

THIS paper presents and analyses data for three linked loci situated at the distal end of the sex chromosome of *Drosophila.*

A systematic analysis was made possible by arrangement of the data in an orthogonal 4×4 Latin square. This arrangement allows factors such as viability and misclassification to be detected and removed so as to obtain accurate estimates of the recombination values. In the past, little use has been made of this design although some examples may be quoted (Anderson, 1921 ; Bridges and Morgan, 1916, 1919, 1923, and Wallace, 1947, 1957).

In the data presented in this paper significant differences of genotypic viability were found. After this viability effect had been allowed for, there was a significant variation in recombination according to the arrangement of the genes in the triple heterozygote. This may be explained as a position effect.

2. METHOD

The loci studied were scute (*sc*), cross-veinless (*cv*) and vermilion (*v*) situated at the distal end of the sex-chromosome. Scute is at 0, *cv* at 13·7, and *v* at 33·0 (Bridges and Brehme, 1944). One Canton-S wild type female was mated to *sc cv v* males and the resultant heterozygous females backcrossed to *sc cv v* males. The segregants from this mating were then intercrossed to obtain the four triply heterozygous females :—

$$\frac{sc\ cv\ v}{+\ +\ +}\ ,\ \frac{sc\ +\ +}{+\ cv\ v}\ ,\ \frac{sc\ cv\ +}{+\ +\ v}\ ,\ \frac{sc\ +\ v}{+\ cv\ +}$$

The members of any one of these four triply heterozygous genotypes were first cousins to the members of the other three. Within the triply heterozygous genotypes $\frac{sc\ cv\ v}{+\ +\ +}$, and $\frac{sc\ cv\ +}{+\ +\ v}$, all females were sibs but within the other two genotypes there were some first cousins. In thc latter case a single mating did not supply enough triple heterozygotes and progeny from a second mating had to be used.

Twenty-five virgin females of each genotype were held until they were 48 hours old to avoid excessive fluctuations in recombination due to age differences (Schultz and Redfield, 1951). They were then mated separately to *sc cv v* males for 24 hours only, to prevent heterogeneity due to overcrowding. The entire experiment was carried out at 25° C. to avoid variations in recombination due to temperature differences (Plough, 1917). When the flies commenced to emerge, classification was carried out daily to minimise any losses after birth due to poor viability of any of the genotypes. Classifying the three mutant genes was easy and could be done

at emergence with the limitation that the wings had to be unfolded to classify cross-veinless.

Four pairs of complementary genotypes are obtained in the progeny :—

sc cv v, + + +	E
+ *cv v*, *sc* + +	F
sc cv +, + + *v*	G
sc + *v*, + *cv* +	H

In relation to any one type of heterozygous female parent, the four pairs of complementary genotypes E, F, G, H, are obtained by four modes of gamete formation. The data are arranged as a 4×4 Latin square in which the columns represent the four modes of gamete formation, and the rows the four types of heterozygous female parents :—

Female parent	Type of gamete formation			
	No change	*sc* change	*v* change	*cv* change
sc cv v/+ + +	E	F	G	H
+ *cv v*/*sc* + +	F	E	H	G
sc cv +/+ + *v*	G	H	E	F
sc + *v*/+ *cv* +	H	G	F	E

TABLE 1

Observed data

Female parent	Type of gamete formation								
	No change		*sc* change		*v* change		*cv* change		
$\frac{sc\ cv\ v}{+\ +\ +}$	*sc cv v*	736	*sc* + +	140	*sc cv* +	206	*sc* + *v*	6	
	+ + +	933	+ *cv v*	103	+ + *v*	274	+ *cv* +	6	
		1669		243		480		12	2404
$\frac{+\ cv\ v}{sc\ +\ +}$	*sc* + +	888	*sc cv v*	143	*sc* + *v*	264	*sc cv* +	8	
	+ *cv v*	758	+ + +	143	+ *cv* +	226	+ + *v*	3	
		1646		286		490		11	2433
$\frac{sc\ cv\ +}{+\ +\ v}$	*sc cv* +	609	*sc* + *v*	162	*sc cv v*	185	*sc* + +	5	
	+ + *v*	741	+ *cv* +	116	+ + +	242	+ *cv v*	10	
		1350		278		427		15	2070
$\frac{sc\ +\ v}{+\ cv\ +}$	*sc* + *v*	825	*sc cv* +	201	*sc* + +	251	*sc cv v*	15	
	+ *cv* +	781	+ + *v*	206	+ *cv v*	234	+ + +	12	
		1606		407		485		27	2525
		6271		1214		1882		65	9432

3. THE EXPERIMENTAL DATA WITH ITS PRELIMINARY ANALYSIS

The data are summarised in the Latin square below (table 1). Each pair of complementary genotypes is split into its components, with the totals shown, these corresponding to the letters of the Latin square.

TABLE 2

Heterogeneity χ^2 values for replicates within each maternal genotype

Maternal genotype	No. of replicates	D.F.	χ^2	Probability
$\frac{sc\ cv\ v}{+\ +\ +}$	25	48	48·4478	0·46
$\frac{sc\ +\ +}{+\ cv\ v}$	21	44	52·4044	0·21
$\frac{sc\ cv\ +}{+\ +\ v}$	20	42	53·8200	0·11
$\frac{sc\ +\ v}{+\ cv\ +}$	25	48	49·0052	0·44

Heterogeneity of the segregations of the replicates within each of the four maternal genotypes was tested (table 2) (Fisher, 1950).

As these χ^2 (table 2) values are not significant, the data can be regarded as homogeneous within each maternal genotype.

Each pair of complementary genotypes is represented once in the offspring of each triply heterozygous female parent. The two members

TABLE 3

Complementary pair	D.F.	Total χ^2	χ^2_1 on total	D.F.	Heterogeneity χ^2
E	4	27·9163 †	26·1523 †	3	1·7646
F	4	18·1636 †	13·4119 †	3	4·7517
G *	3	21·5455 †	17·7963 †	2	3·7519
H	4	11·7654 ‡	6·8667 †	3	4·8987

* In G the double recombinant class was added to one of the single recombinant classes as it was very small.

† Significant at P = 0·01.

‡ Significant at P = 0·02.

of the pair, assuming no viability or other disturbance, should be in a 1 : 1 ratio. Assuming a 1 : 1 ratio, a χ^2 test may be carried out for each pair of complementary genotypes for each of the four heterozygous female parents separately, and also on the grand totals of each complementary pair. The difference between the sum of the four χ^2 values for one degree of freedom each, and the χ^2_1 on the grand total will give an approximate heterogeneity χ^2 for each complementary pair (table 3).

From table 3 it can be concluded that members of each complementary pair are not in a 1 : 1 ratio and that deviations from the 1 : 1 ratio are homogeneous. This result suggests a consistent viability disturbance in the data.

From the marginal totals of the observed data (table 1) proportionate expectations were calculated for each cell of the 4×4 table (table 4).

TABLE 4

Expected frequencies

1598·3338	309·4207	479·6785	16·5670	2404·0000
1617·6148	313·1533	485·4650	16·7669	2433·0000
1376·2691	266·4313	413·0344	14·2653	2070·0000
1678·7823	324·9947	503·8221	17·4008	2525·0000
6271·0000	1214·0000	1882·0000	65·0000	9432·0000

Comparison of these expectations with the observed data (table 1) gives a general test of disturbance. In this case χ^2 was 54·8798 for nine degrees of freedom which is highly significant.

Viability differences between the pairs of genotypes E, F, G, H, which may have disturbed the linkage relationships will become evident by comparison of observed with the expected frequencies (Fisher, 1949), thus :

Complementary pair	Expected	Observed	χ^2
E	2341·9223	2409	1·9212
F	2445·1229	2389	1·2882
G	2197·7092	2248	1·1508
H	2447·2456	2386	1·5328
	9432·0000	9432	

giving $\chi^2 = 5{\cdot}8930$ for three degrees of freedom such that $0{\cdot}1 < P < 0{\cdot}2$. Using Fisher's (1949) method a more exact χ^2_3 was 12·6078 which is significant at $P = 0{\cdot}01$. This viability χ^2_3 is, however, a small portion of the total $\chi^2_9 = 54{\cdot}8798$.

4. REMOVAL OF VIABILITY DISTURBANCE—METHOD

As a first approximation, the method of Fisher (1949) is appropriate to remove viability disturbances, but the following iterative method due to Dr A. R. G. Owen provides a solution to any desired degree of accuracy.

Let p, q, r, s be the frequencies of the four modes of gamete formation and a, b, c, d be unknowns proportional to the number of zygotes found in each mating type (heterozygous female parent). Let u, v, w, t, be the average viability of the complementary pairs E. F, G, H.

The definitions can be standardised by :

$$\begin{aligned} u+v+w+t &= 4 \\ a+b+c+d &= 1 \\ p+q+r+s &= 1 \end{aligned}$$

Expectations can now be drawn up :

$$n\mathrm{K}\times \left\{ \begin{array}{llll} apu & aqv & arw & ast \\ bpv & bqu & brt & bsw \\ cpw & cqt & cru & csv \\ dpt & dqw & drv & dsu \end{array} \right.$$

such that n = number bred, and $\lambda = \frac{1}{\mathrm{K}} = apu+\ldots cqt\ldots+dsu$, 16 terms in all.

Let the observations divided by n be

a_1	a_2	a_3	a_4	A
b_1	b_2	b_3	b_4	B
c_1	c_2	c_3	c_4	C
d_1	d_2	d_3	d_4	D
P	Q	R	S	1

where row and column totals are as shown and U, V, W, T, are the totals of the pairs E, F, G, H.

The following transformations can be made :

$$\begin{aligned} x_1 &= a+b+c+d & 4a &= x_1+x_2+x_3+x_4 \\ x_2 &= a-b+c-d & 4b &= x_1-x_2+x_3-x_4 \\ x_3 &= a+b-c-d & 4c &= x_1+x_2-x_3-x_4 \\ x_4 &= a-b-c+d & 4d &= x_1-x_2-x_3+x_4 \end{aligned}$$

$$\begin{aligned} y_1 &= p+q+r+s & 4p &= y_1+y_2+y_3+y_4 \\ y_2 &= p-q+r-s & 4q &= y_1-y_2+y_3-y_4 \\ y_3 &= p+q-r-s & 4r &= y_1+y_2-y_3-y_4 \\ y_4 &= p-q-r+s & 4s &= y_1-y_2-y_3+y_4 \end{aligned}$$

$$\begin{aligned} 4z_1 &= u+v+w+t & u &= z_1+z_2+z_3+z_4 \\ 4z_2 &= u-v+w-t & v &= z_1-z_2+z_3-z_4 \\ 4z_3 &= u+v-w-t & w &= z_1+z_2-z_3-z_4 \\ 4z_4 &= u-v-w+t & t &= z_1-z_2-z_3+z_4 \end{aligned}$$

and let $\mathrm{X}_1 = \mathrm{A}+\mathrm{B}+\mathrm{C}+\mathrm{D}$ $\quad \mathrm{Y}_1 = \mathrm{P}+\mathrm{Q}+\mathrm{R}+\mathrm{S}$

$\mathrm{X}_2 = \mathrm{A}-\mathrm{B}+\mathrm{C}-\mathrm{D}$ etc. $\quad \mathrm{Y}_2 = \mathrm{P}-\mathrm{Q}+\mathrm{R}-\mathrm{S}$ etc.

$$\begin{aligned} 4\mathrm{Z}_1 &= \mathrm{U}+\mathrm{V}+\mathrm{W}+\mathrm{T} \\ 4\mathrm{Z}_2 &= \mathrm{U}-\mathrm{V}+\mathrm{W}-\mathrm{T} \text{ etc.} \end{aligned}$$

Putting $x_1 = \mathrm{X}_1 = y_1 = \mathrm{Y}_1 = z_1 = \mathrm{Z}_1 = 1$ will give a set of

equations which by iteration will give a maximum likelihood solution to any desired degree of accuracy.

$$\begin{aligned}
\lambda &= 1+x_2y_2z_2+x_3y_3z_3+x_4y_4z_4\\
x_2 &= \lambda X_2-y_2z_2-x_3y_4z_4-x_4y_3z_3\\
x_3 &= \lambda X_3-y_3z_3-x_2y_4z_4-x_4y_2z_2\\
x_4 &= \lambda X_4-y_4z_4-x_2y_3z_3-x_3y_2z_2\\
\\
y_2 &= \lambda Y_2-x_2z_2-y_3x_4z_4-y_4x_3z_3\\
y_3 &= \lambda Y_3-x_3z_3-y_2x_4z_4-y_4x_2z_2\\
y_4 &= \lambda Y_4-x_4z_4-y_2x_3z_3-y_3x_2z_2\\
\\
z_2 &= \lambda Z_2-x_2y_2-z_3x_4y_4-z_4x_3y_3\\
z_3 &= \lambda Z_3-x_3y_3-z_2x_4y_4-z_4x_2y_2\\
z_4 &= \lambda Z_4-x_4y_4-z_2x_3y_3-z_3x_2y_2
\end{aligned}$$

From these, values for a, b, c, d and p, q, r, s and u, v, w, t can be obtained. The values p, q, r, s will give the frequencies of modes of gamete formation from which, by addition, the recombination values may be calculated.

5. REMOVAL OF VIABILITY DISTURBANCE FROM THE DATA UNDER DISCUSSION

The method outlined in the last section gave values for p, q, r, s :

$$\begin{aligned}
p &= 0{\cdot}665{,}685\\
q &= 0{\cdot}127{,}645\\
r &= 0{\cdot}199{,}835\\
s &= 0{\cdot}006{,}835
\end{aligned}$$

From these, the corrected recombination values shown in table 5 were calculated. These are compared with the crude values calculated directly from the observed data.

TABLE 5

Crude and corrected recombination values

	sc-cv	*cv-v*	*sc-v*
Crude	13·5602	20·6425	32·8244
Corrected	13·4480	20·6670	32·7480

The corrected values are not much different from the crude values. Corrected values calculated using Fisher's (1949) method, which is a first approximation only, gave values intermediate between the crude and corrected values in table 5 thus :

sc—cv	*cv—v*	*sc—v*
13·4965	20·6583	32·7828

A reconstructed table of expectations using the iterative solutions for p, q, r, s ; a, b, c, d, and u, v, w, t may now be drawn up (table 6). The entries in this table are the expected frequencies after the elimination of viability.

A useful check is easily made by seeing if the rows, columns, and totals E, F, G, H are the same in table 6 as before iteration. This

TABLE 6

Expected frequencies (viability removed)

1625·9907	279·6249	483·4746	14·8838	2403·9740
1596·9582	341·4333	476·5399	18·1089	2433·0403
1393·9628	240·5798	422·4752	12·9596	2069·9974
1654·0591	352·3816	499·5176	19·0500	2525·0083
6270·9708	1214·0196	1882·0073	65·0023	9432·0000

check is a direct consequence of the maximum likelihood method used in the elimination of viability (Section 4).

Comparing the entries in table 6 with the observed values (table 1), gave a $\chi^2 = 41{\cdot}3924$ for six degrees of freedom. Three degrees of freedom were lost in the estimation and removal of viability interaction.

Thus we have :

DF	χ^2	
9	54·8798	Comparing expected and observed frequencies.
6	41·3924	Comparing expected and observed frequencies after removing viability.
3	13·4874	Viability.

TABLE 7

Maternal genotype	Recombination fractions		
	sc-cv	*cv-v*	*sc-v*
(i) *sc cv v*/+ + + .	10·6073±0·6199	20·4659±0·8229	30·0749±0·9353
(ii) *sc* + +/+ *cv v* .	12·2072±0·6637	20·5919±0·8199	31·8948±0·9449
(iii) *sc cv* +/+ + *v* .	14·1546±0·7662	21·3527±0·9007	34·0580±1·0416
(iv) *sc* + *v*/+ *cv* + .	17·1881±0·7508	20·2772±0·8001	35·3267±0·9512

In an attempt to detect factors responsible for this highly significant disturbance $\chi^2_6 = 41{\cdot}3924$ recombination values for each maternal genotype were calculated (table 7). The recombination values for the segment *sc—cv* were found to vary significantly according to the maternal genotype but in the segment *cv—v* there is no such variation. The variable recombination values for the entire segment *sc—v* reflect the variation in segment *sc—cv*. From examination of table 7 it will be evident that the data from each of the four maternal genotypes should be treated separately.

In this experiment environmental and genetic factors likely to cause variation were controlled. That such measures were effective can be seen from the homogeneity of families within each mating type (table 2). If a lethal or structural re-arrangement had affected part of the experiment it would be expected that the heterogeneity χ^2 values (table 3) for comparisons within each complementary pair would be significant and probably the values in table 2 also. However, if such a phenomenon had affected all the members of one female genotype equally the χ^2 values for table 2 would be expected to be non-significant. It can be concluded from the lack of heterogeneity in both of these tests that there was no lethal or structural re-arrangement affecting the data.

Dubinin (1933) discusses step-allelomorphism of the achaete-scute gene which would lead to misclassification if present. The scute gene that was used was not of variable manifestation and there was no indication of any error in breeding due to misclassification.

TABLE 8

Kosambi coefficients

Maternal genotype	K	χ^2_1 for deviation from unity
sc cv v/+ + + . .	0·3800±0·1011	4·5502
sc + +/+ *cv v* . .	0·2820±0·0854	5·6480
sc cv +/+ + *v* . .	0·3512±0·0903	6·4076
sc + *v*/+ *cv* + . .	0·4342±0·0833	8·7811
Whole experiment . .	0·3750±0·0464	25·7191

6. KOSAMBI COEFFICIENTS

Kosambi coefficients (Owen, 1953) may be calculated using the formula :

$$K = \frac{n^2 n_{12}}{2n_1 n_2 n_{1+2}}$$

where $n_1 = a_1 + a_{12}$
$n_2 = a_2 + a_{12}$
$n_{1+2} = a_1 + a_2$
$n_{12} = a_{12}$

and a_1 show recombination in segment 1 (sc—cv)
a_2 show recombination in segment 2 (cv—v)
a_{12} show simultaneous recombination in both segments.

To test departure of $\hat{K}$ from unity we may calculate

$$\chi^2_1 = \frac{(\hat{K}-1)^2}{V(K)}$$

where V(K) is the sampling variance at K = 1 (Owen, 1953). These tests are tabulated in table 8.

Despite the variation in the recombination fraction according to

maternal genotype, the four K values, which are all significantly less than unity (table 8), do not differ significantly from each other. Low K values such as these are expected for the distal end of the long arm of a chromosome on the basis of Owen's theory of genetic recombination (Owen, 1950).

7. DISCUSSION

A brief mention of the earlier use of orthogonal design in three-point data is appropriate. Bridges and Morgan (1916, 1919, 1923) published three-point data in *Drosophila* using this design. They assumed that tabulation of their data into a 4×4 Latin square would in itself remove spurious effects without further calculation. If only one gene is of lowered viability, calculation of recombination values from the column totals of the observed data is accurate, but if more than one gene is of lowered viability preliminary calculation of the type described by Fisher (1949) or the method outlined in this paper must be employed. There were no tests to detect differential viability or any other type of disturbance in these early data. It cannot be ascertained to what degree the environment was controlled in these experiments but there is frequently heterogeneity between individual matings and in some cases heterogeneity between members of a pair of complementary genotypes. In maize, Anderson (1921) published data arranged as a 4×4 Latin square, but they are heterogeneous probably due to misclassification which was mentioned as a source of error by the author.

More recently Wallace (1947, 1957) has published homogeneous data for the house mouse. In the former paper recombination values significantly greater than 50 per cent. were detected in the sex chromosome and the latter paper refers to an experiment using three markers of linkage group V to obtain accurate interference data.

It can be concluded that much of the early three-point linkage data is not of great accuracy, either because of heterogeneity, or because of some disturbance, undetected due to the lack of orthogonality. Wallace's recent work on mice is, of course, excluded from this criticism for it shows how these difficulties may be overcome.

We have at present little knowledge as to variation produced in the estimate of recombination value caused by extraneous genetic and environmental factors. In the experiment described in this paper, factors likely to cause heterogeneity were controlled as much as possible.

The orthogonal 4×4 Latin square design used in the analysis enabled viability disturbances to be removed by the method given in this paper, so as to obtain good estimates of the recombination values. This design is most efficient when the total progeny from each of the four mating types are approximately equal. In the data under discussion totals were 2404, 2433, 2070 and 2525 giving tolerable efficiency.

As well as a viability disturbance in the data, there was found to be a significant variation in recombination according to the arrangement of the genes in the maternal genotype. This variation occurs in the distal segment of the chromosome only, namely *sc—cv*, and not in the segment *cv—v*. The variation in this segment cannot be ascribed to the coupling and repulsion arrangement of the genes in the parent, since, as can be seen from table 8, maternal genotypes (i) and (iii) are in coupling for *sc—cv*, and (ii) and (iv) in repulsion, and the members of each pair differ greatly. The genes *cv—v* are in coupling in (i) and (ii) and in repulsion in (iii) and (iv) ; the members of each pair are unequal and are only sub-significantly different. Even so, the recombination value varies sufficiently to regard the offspring from the four different maternal genotypes as separate entities.

This phenomenon is in all probability due to a type of position effect dependant upon the arrangement of the genes in the heterozygote. Such a position effect may, of course, be restricted to the distal end of the chromosome. Equally careful tests in other regions of the same and different chromosomes should throw light on this interpretation.

8. SUMMARY

1. An orthogonal three-point linkage test was carried out in the sex chromosome of *Drosophila melanogaster*.

2. Significant differences in genotypic viabilities were found. Allowance was made for these differences in estimating recombination values for the three segments.

3. The estimates of the recombination values so obtained were found to vary according to the arrangement of the genes in the heterozygous parent. This may be interpreted as a position effect.

4. Kosambi values were calculated for each maternal genotype. These were found to be homogeneous and all significantly less than unity.

5. A brief mention is made of adequate control and orthogonal design in three-point linkage studies. Without these precautions the position effect discussed could not have been detected.

Acknowledgments.—I wish to thank Professor D. G. Catcheside, F.R.S., of the Genetics Department, University of Adelaide, who suggested this problem, Professor Sir Ronald Fisher, F.R.S., for helpful discussion on the statistical matters, and Dr A. R. G. Owen for the iterative method of removing viability disturbances.

9. REFERENCES

ANDERSON, E. G. 1921. The inheritance of salmon silk color in maize. *Cornell Univ. Agric. Exper. Station, Memoir 48*, 535-554.

BRIDGES, C. B., AND BREHME, K. S. 1944. The mutants of *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.*, p. 552.

BRIDGES, C. B., AND MORGAN, T. H. 1916. Sex-linked inheritance in *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.*, p. 237.

BRIDGES, C. B., AND MORGAN, T. H. 1923. The third chromosome group of mutant characters of *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.*, p. 327.

DUBININ, N. P. 1933. Step-allelomorphism in *Drosophila melanogaster*. *J. Genet.*, *37*, 443-464.

FISHER, R. A. 1949. Note on the test of significance for differential viability in frequency data from a complete three point test. *Heredity*, *3*, 215-219.

FISHER, R. A. 1950. *Statistical Methods for Research Workers*. 11th Ed. Oliver and Boyd.

MORGAN, T. H., AND BRIDGES, C. B. 1919. The second chromosome group of mutant characters of *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.*, p. 278.

OWEN, A. R. G. 1950. The theory of genetical recombination. *Advances in Genetics*, *3*, 117-157.

OWEN, A. R. G. 1953. The analysis of multiple linkage data. *Heredity*, *7*, 247-264.

PLOUGH, H. H. 1917. The effect of temperature on crossing over in *Drosophila*. *J. Exp. Zool.*, *24*, 147-209.

SCHULTZ, J., AND REDFIELD, H. 1951. Interchromosomal effect on crossing over in *Drosophila*. *Cold Spring Harbor Symposia on Quantitative Biology*, *16*, 175-197.

WRIGHT (= WALLACE), M. E. 1947. Two sex-linkages in the house mouse with unusual recombination values. *Heredity*, *1*, 349-354.

WALLACE, M. E. 1957. *Heredity* (in press).