

## GENETICS OF COXAL CHAETAE IN *DROSOPHILA MELANOGASTER*

### I. VARIATION IN GENE ACTION

KENNETH MATHER and M. J. HANKS

*Department of Genetics, University of Birmingham, Birmingham B15 2TT*

Received 29.iv.77

#### SUMMARY

Differences observed between the numbers of chaetae on the coxae of corresponding legs on the two sides of a fly are most probably due to developmental instability, as with sternopleurals. The metric used in the present study has therefore been taken as the sum of the chaetae on the coxae of the pair of corresponding legs. The average number is lowest on the rear (R), highest on the front (F) and intermediate on the middle (M) legs. The number of coxals on the M legs often approximates to the average of R and F in other cases being most commonly higher than this average: indeed at 25°C no case has been observed of M falling significantly below the average of R and F. A reduction of temperature to 18°C almost always raises the number of chaetae on the R legs, but usually, though not always, it reduces the numbers on the F legs. The number of chaetae on the M legs is always lower at 18°C than at 25°C except for the case of Australia 8.

Chromosome assays based on sets of substitution lines made between the Wellington (W), Samarkand (S) and 6C/L (L) inbred lines, taken in pairs, show all the major chromosomes to be carrying genes affecting the numbers of coxal chaetae. A further analysis of the X chromosome using a marked tester chromosome (T) reveals at least four and possibly as many as seven loci affecting coxal chaeta number, with a concentration at the left end of the chromosome.

Regression analysis of the type used in investigating genotype  $\times$  environment interactions shows that the rise in chaeta number from R through M to F cannot be due to a uniformly proportionate enhancement of the effects of the genes concerned, and this is confirmed by a closer analysis of the differences between the homologous chromosomes revealed by the chromosome assay: a chromosome which gives a higher number of chaetae than its homologue on one pair of legs can give a lower number on a different pair. The genes must in fact show differential changes of activity from one pair of legs to another. The results can be accounted for by postulating three classes of gene: one class ( $\alpha$ ) being equally active on all three pairs of legs; a second ( $\beta$ ) inactive on R but active on both M and F; and a third class ( $\gamma$ ) inactive on R and M but active on F. This hypothesis allows predictions to be made about the responses of coxal chaeta number to certain types of selection, the results of which will be the subject of a later publication.

The differential changes in the action of the genes could perhaps be understood in terms of a uniform activity of the genes themselves being modified by appropriate variation in the availability in the cell of other materials necessary for the gene products to display their activity by chaeta production. A more attractive interpretation, however, is to be found in terms of regulatory, control and structural elements similar to the tripartite system postulated and discussed by Britten and Davidson.

#### I. INTRODUCTION: THE PROBLEM

THE coxa of each of the six legs in *Drosophila melanogaster* carries a group of chaetae, the number in the group varying between different legs of the same

fly as well as between corresponding legs of different flies. The variation is quasi-continuous just like that of the more familiar sternopleural chaetae; but unlike the sternopleurals, little attention has so far been paid to the coxal chaetae. Mather and Harrison (1949) recorded a correlated response in the number of coxal chaetae on the rear legs of *Drosophila* to selection for the number of abdominal chaetae, and Spickett (1963) observed that a line of flies carrying a factor for a high number of sternopleurals also had an increased number of coxals on the front and middle legs. Beyond this, however, little if any genetical information appears to be available about coxal chaeta number.

The presence of three pairs of legs, front (F), middle (M) and rear (R) allows two kinds of comparison to be made and two kinds of information to be obtained. Comparisons between the numbers of chaetae borne by the coxae of the two legs, left and right, of a pair gives essentially the same type of information as has been obtained from similar comparisons of sternopleurals on the left and right side of the fly. As with sternopleurals we have found no evidence of lateral asymmetry such as would be exemplified by a difference in the mean numbers of coxals on the two sides. Rather the differences between sides, which favour one side as often as the other, stem from instability in development, though the magnitude of the differences, and hence of the degree of instability which they exemplify, would be expected to be subject to genetical control as has been found for the sternopleurals (see Mather, 1953). This type of comparison will be pursued no further, and the numbers of coxal chaetae used in this report will be the sum of the numbers on the members of the pair of legs in question, in the same way that the number of sternopleurals is commonly given as the sum of the numbers on the two sides.

The second kind of comparison is between the three different pairs of legs and this allows us to obtain information about differences which arise as part of the antero-posterior differentiation of the fly, and also about its genetical control. Though corresponding parts of the F, M and R legs must be basically homologous to one another, their coxae are not alike in shape, as will be seen from fig. 1, which illustrates the relevant parts of the right F, M and R legs of a single fly. Nor are the numbers of coxal chaetae that they bear alike, though the differences in their numbers are not simply relatable to the differences in shape.

The number of chaetae borne by the coxae of the F, M and R pairs of legs in each of eight inbred lines of flies are shown graphically in fig. 2. The coxals were counted on each of five males and five females from each line raised at 25°C and 18°C, and the mean numbers for each pair of legs are shown in the fig. The eight lines illustrated are taken from a wider sample of 24 inbred lines that were counted of which 18 were derived from the Texas population maintained in this laboratory. They were chosen to illustrate the variation that was found in coxal chaeta numbers. Being chosen for this purpose they illustrate *inter alia* the extremes of both overall numbers and contrast between the three pairs of legs: the remaining lines fall within the range illustrated.

A number of points emerge clearly from these illustrations. First the number of coxal chaetae always rises from the rear to front legs. Even in Texas 9 at 25° where the M chaetae are nearly as numerous on average as the F, the F mean is 22·15 while that of M is 22·00. Secondly, while the

average number of chaetae on the M coxae often approximates closely to the mid-point between F and R, in some cases it departs significantly from this point, the most striking example again being provided by Texas 9. At 25° this departure is always in the direction of the M mean being higher than the mid-value of F and R, apart from Texas 25 where it is marginally (and not significantly) lower than this mid-value. At 18°, however, there is better

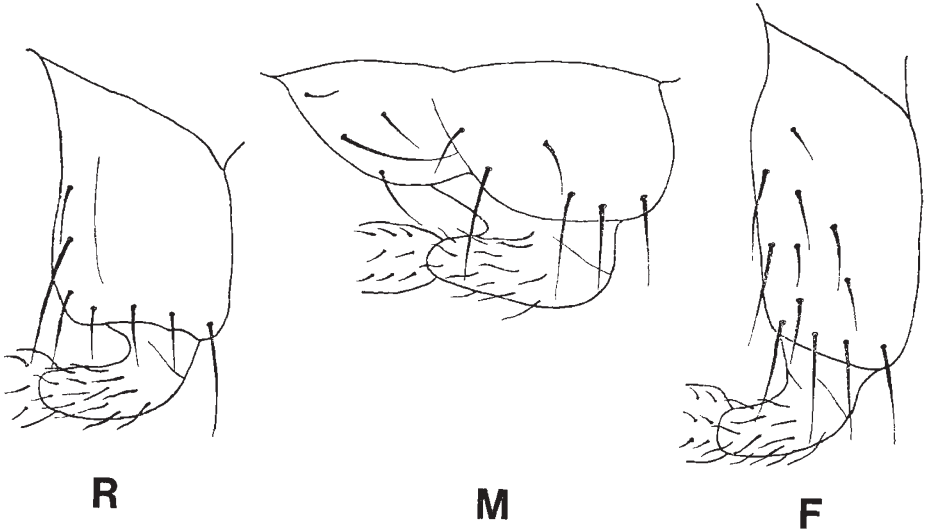


FIG. 1.—Sketches showing the coxae of the rear (R), middle (M) and front (F) legs from the right side of a female. The three coxae carry 7, 10 and 12 chaetae respectively. The chaeta numbers used in the text are the sums of the numbers on the corresponding right and left legs.

evidence that the number of M coxals can fall below the mid-value, as is seen for example in Texas 8. There are thus differences between the legs in their responses to change in temperature. If the R chaeta number shows a significant response to change in temperature it is by being higher at 18° than at 25°. The F and M coxals most often respond in the other direction, giving a higher number at 25° than at 18° though Australia 8 affords a clear exception the numbers being higher at 18° than at 25°.

The final point is that there are clearly genetical differences between the lines both in the overall numbers of coxal chaetae and in the pattern of rise of their number from R through M to F. Comparisons among the Texas lines show that these genetical differences in both overall number and pattern of rise can exist within a single population. Furthermore, the different lines respond differently to temperature change, so showing that genotype  $\times$  environment interaction is involved in the determination of coxal chaeta number.

These observations raise two questions. What is the nature of the genetical differences between the lines? And how do the genes act to produce the rise in the number of coxal chaetae from the rear to the front legs, which as we have seen varies in the pattern of its progress from R to M to F, as well as in its magnitude? In addressing ourselves to these

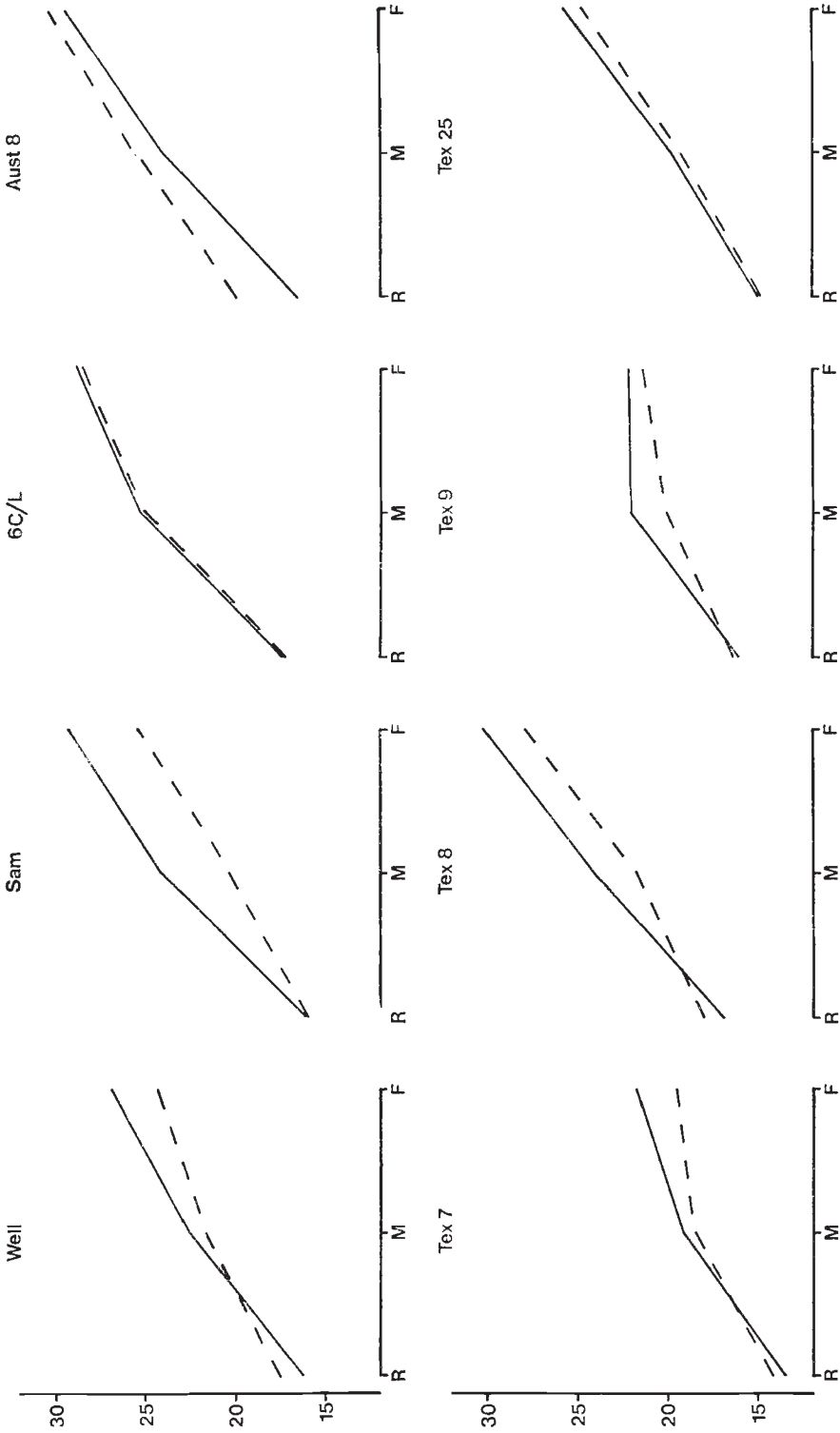


Fig. 2.—The average numbers of chaetae on the coxae of the rear (R), middle (M) and front (F) pairs of legs of flies from eight inbred lines. Well, Sam and 6C/L are the inbreds of which substitution sets of lines were made, as described in the text. Aust 8 is an inbred derived from an Australian population, and Tex 7, 8, 9 and 25 are four inbreds from the Texas population. The eight lines illustrate the range in numbers of chaetae, patterns of change from leg to leg, and response to change in temperature. The solid lines show the numbers of chaetae borne by flies raised at 25°C and the broken lines the numbers borne

questions, we would emphasise that all the flies we are discussing were wild-type, and the variation was such as to be found in wild populations—indeed the differences among the Texas inbred lines must have been present in the population from which they were derived. It is in fact variation which must have been subject to, and have survived, the impact of natural selection in the past.

## 2. THE GENETICAL DIFFERENCES BETWEEN LINES

### (i) *Chromosome assays*

The nature of the genetical differences between lines was investigated by the use of three sets of substitution lines between the Wellington, Samarkand and 6C/L inbreds. A set of substitution lines consists of the eight true-breeding lines comprising all the possible combinations of the three major chromosomes taken as units from two parental lines. Thus the set W/S between Wellington and Samarkand consists of the true-breeding lines WWW, WWS, WSW, WSS, SWW, SWS, SSW and SSS where W and S indicate Wellington and Samarkand chromosomes respectively, with the X, II and III chromosomes written in that order. Such a set of lines involving W and S was already available from the experiments on genotype  $\times$  environment interaction described by Caligari and Mather (1975). Similar substitution sets were made up between Samarkand and 6C/L (S/L) and between 6C/L and Wellington (L/W). They of course comprised respectively the chromosome combinations SSS, SSL, SLS, SLL, LSS, LSL, LLS, LLL and LLL, LLW, LWL, LWW, WLL, WLW, WWL, WWW.

The method by which the W/S set was constructed is described by Caligari and Mather (*loc. cit.*). The other two sets were made in a somewhat simpler way suggested by our colleague Dr M. J. Kearsey which does not involve the use of a marked X chromosome carrying an inversion, and requires fewer generations than that used by Caligari and Mather. The marked II and III chromosomes were the same as those used by Caligari and Mather. Only one substitution set was made between Samarkand and 6C/L, but two sets were made independently of each other between 6C/L and Wellington, one of them by Dr P. D. S. Caligari to whom we are indebted for its use.

The numbers of coxal chaetae were determined for each substitution set at two temperatures, 18° and 25°C. The three substitution sets were raised for this purpose in separate experiments carried out on different occasions. Comparisons between substitution sets are thus subject to a standard error potentially greater than that of comparisons among lines from the same set. Each culture was the product of eight females mated to 10 males and allowed to lay for five days in a  $\frac{1}{3}$  pint milk bottle containing standard yeasted food. The chaetae on all six legs were counted for 10 flies of each sex and the means from the two sexes were taken as the basic data from the culture. (As already noted, the number of coxals from corresponding legs on the two sides were added together.) Each line was raised in duplicate. Thus the experiment using the W/S set of lines included 32 cultures, *viz.* two for each of the eight lines at each of two temperatures. The experiment with the S/L set was the same, but that with L<sub>r</sub>/W was twice as large, comprising 64 cultures, since there were two independent sets of lines for this combination.

A summary of the results from all three substitution sets is given in

table 1, and a summarised analysis of variance of the results in table 2. Taking table 2 first, there are 64 basic observations in the W/S set, *viz.* the two sex means from each of 32 cultures. Of the 63 degrees of freedom (d.f.) 32 are between corresponding observations on the two sexes in the duplicate cultures from the eight lines at the two temperatures. This yields the item for replicate error variation in the table. The remaining 31 d.f. can be partitioned into 7 among the eight lines, 1 between the two temperatures, 1 between the two sexes, 7 for interaction between lines and temperatures ( $C \times T$ ), 1 for interaction between temperatures and sexes ( $T \times S$ ), 7 for interaction between lines and sexes ( $C \times S$ ) and lastly 7 for the composite interaction between lines, temperatures and sexes ( $C \times T \times S$ ). The d.f. among lines can of course be partitioned further into 1 each for the main effects of the three chromosomes, X, II and III, and 4 for interactions among them. These 4 could have been broken down further, but are in fact presented as a pool.

The analysis of the S/L set of lines takes exactly the same form as for W/S; but the L/W set differs in that there are twice as many cultures and hence twice as many observations, *viz.* 128 in all, there being two parallel substitution sets duplicated in the experiment. Thus each sex of each of the eight lines in the set contributes not 1 but 3 d.f. to the error variance. Two of these, between the duplicates of the independent but corresponding genotypes, are replicate error, which thus has  $2 \times 32 = 64$  d.f. The third is between corresponding genotypes in the two substitution sets and will reflect the effects of residual recombination, evidence of which Caligari and Mather (*loc. cit.*) found in their experiments, or indeed any other effect such as mutation by which the corresponding genotypes may have come to differ. This accounts for the remaining 32 d.f. in the analysis. The 31 d.f. relating to the effects of lines, temperature and sexes are partitioned just as in the W/S and S/L analyses.

The recombination error variance for 32 d.f. is not shown in table 2: it showed evidence of significance especially in the data from the F legs when tested against the replicate error, but since we have no counterpart for it in W/S and S/L a different estimate of error variance was used in all the analyses. This was obtained from the pool of the  $C \times S$ ,  $T \times S$  and  $C \times T \times S$  interactions which jointly correspond to 15 d.f. When tested against the Replicate Error, the mean square for these pooled interactions shows significance for the F legs in two analyses and although not formally significant is higher than the Replicate Error in five of the six tests on the M and R legs (table 2). It has therefore been used as the Pooled Error for the next test of significance, which was that of the  $C \times T$  interaction. Six of the nine tests of  $C \times T$  against the Pooled Error prove to be significant. The  $C \times T$  mean square must thus be used as the estimate of error for testing the mean squares among the eight lines.

Each of the nine  $C \times T$  mean squares has only 7 d.f. and tests of significance based on them individually will thus not be very sensitive. We may note therefore that while Bartlett's Test shows these  $C \times T$  mean squares to be heterogeneous over legs, it also shows them to be homogeneous over substitution sets. We may therefore pool over substitution sets the  $C \times T$  interactions shown by a given pair of legs, to obtain a combined mean square for 21 d.f. The combined mean squares so obtained are 2.469, 1.987 and 0.794 for the F, M and R legs respectively. Using these mean squares we

TABLE I  
Mean numbers of coxal chaetae in the three sets of substitution lines at 18°C and 25°C

	18°C			25°C			18°C			25°C		
	F	M	R	F	M	R	F	M	R	F	M	R
WWW	24.35	21.35	18.05	26.00	22.60	16.55	28.10	21.85	16.35	30.10	24.40	16.15
WWS	24.60	21.00	17.80	27.10	22.25	16.95	28.35	21.25	16.30	30.00	23.50	15.90
WSW	26.25	21.85	17.75	27.60	22.35	16.50	25.00	23.30	17.65	25.65	24.20	17.50
WSS	24.00	21.10	16.40	28.10	22.95	16.65	24.30	22.90	16.65	24.40	24.65	16.70
SWW	25.25	19.90	15.50	26.55	22.70	15.70	29.90	21.80	16.90	29.15	24.55	16.25
SWS	25.35	19.50	15.50	27.45	22.20	15.90	29.45	22.20	16.55	29.50	23.40	16.35
SSW	25.60	20.70	17.50	27.25	22.35	15.25	26.55	21.25	17.85	25.95	23.90	17.55
SSS	24.85	20.10	16.55	26.10	21.50	15.90	25.40	23.65	16.75	26.05	23.90	16.95
	26.10	20.20	16.50	26.20	21.05	16.30	30.80	23.00	16.75	33.75	26.35	16.00
	24.80	19.00	15.80	29.75	23.35	15.80	31.05	23.00	16.80	33.20	25.00	15.95
	23.75	19.65	15.50	29.35	23.05	16.15	28.75	25.25	17.80	27.90	24.95	17.40
	22.15	19.20	14.85	26.60	21.65	15.90	28.25	24.65	18.00	26.90	24.95	17.20
	27.55	20.90	16.75	30.70	25.05	17.75	29.80	25.10	17.50	30.75	24.30	17.25
	26.45	20.55	16.20	30.85	24.65	17.00	28.95	25.00	17.70	28.80	25.75	16.60
							28.25	24.85	18.25	28.80	26.20	17.20

The top figure of each pair is the male mean and the lower figure the female mean.

TABLE 2  
*Analysis of variance of chaeta numbers in the three sets of substitution lines*

Item	d.f.	Mean squares											
		F	M	R	F	M	R	F	M	R	F	M	R
X	1	5.523	1.960	2.031	103.276***	43.725**	2.764	422.678***	126.803***	10.294**			
II	1	0.040	0.040	11.223***	0.701	0.926	1.568	81.920***	3.380	25.472***			
III	1	64.000***	6.631	0.023	199.869***	8.194*	10.160**	93.503***	53.561***	9.299**			
Inls.	4	7.715*	6.106*	3.648**	4.425	0.334	0.531	0.881	1.250	0.639			
T-Temps	1	105.063	69.723	2.806	5.581	22.920	3.106	93.845	22.781	16.893			
S-Sexes	1	0.076	1.563	0.063	3.469	0.926	0.473	6.480	0.025	0.310			
C x T	7	2.475*	3.192***	1.041**	2.586***	2.376**	0.205	2.346	0.394	1.137***			
Pooled Ints.†	15	0.634*	0.145	0.203	0.303	0.393	0.236	0.932**	0.280	0.237			
Replicate errors	32§	0.275	0.162	0.163	0.648	0.272	0.174	0.312	0.190	0.174			
			W/S			S/L			L/W				

† Pooled Interactions is pool of C x S, T x S and C x T x S, with d.f. = 7, 1 and 7 respectively.

§ d.f. = 64 for L/W.

P = \*0.05-0.01; \*\* P = 0.01-0.001; \*\*\* P < 0.001.

C Items are tested against C x T, C x T against Pooled Interactions and Pooled Ints. against Replicate Error.



have tested the contribution of the three chromosomes to the variation among the lines of each substitution set, with the results set out in table 2.

There is evidence of interaction among the chromosomes in the W/S set, but none in the S/L and L/W sets. Turning to the main effects of the chromosomes, we find that all three of them are implicated. The W/S set implicate both chromosomes II and III in the production of the differences between Wellington and Samarkand, while the S/L set implicates X and III. The L/W set reveals significant effects for all three chromosomes. We note, however, that the chromosomes do not have corresponding effects on the chaeta number of all three pairs of legs. Thus in W/S only the F legs reveal a significant effect of chromosome III but only the R legs a significant effect of chromosome II. Similarly in S/L, R shows no effect of the X, though both M and F display one. Even in L/W no significant effect of chromosome II is displayed by M, despite the clear effects of this chromosome on F and R. These results will be discussed further in Section 3. In the meantime we note that all three major chromosomes must carry genes mediating the numbers of coxal chaetae, though they appear not to be affecting the different legs in the same way.

#### (ii) *The X chromosome*

In the previous section the three chromosomes were followed as units in their effects on coxal chaeta number. The overall effect of a chromosome is, however, the sum of the effects of the individual genes that it carries and the effect of the chromosome may even be undetectable where several gene differences are present if they are balancing one another's effects by acting in opposite directions. A further analysis of the X chromosome has therefore been undertaken to see whether there was evidence of genes at more than one locus being involved in the production of the effects that the chromosome assays revealed it as displaying.

This analysis was made using the method described by Wigan (1949). An X chromosome (T) marked by the genes  $y^2w^a$  cv ct v f was introduced into the WWW line, so giving a new line TWW. Females of lines WWW, SWW and LWW were mated to TWW males. Daughters from these three crosses (heterozygous respectively for the T/W, T/S and T/L X chromosomes but all homozygous WW for chromosomes II and III) were mated to WWW males and allowed to lay in  $\frac{1}{8}$  pint milk bottles. Three replicate cultures were raised from each of the three  $F_1$ 's, thus making nine cultures in the whole experiment.

The male offspring from these cultures showed of course segregation for the marker genes. The numbers of coxal chaetae on the three pairs of legs were counted on 10 males each of the two parental classes,  $y^2w^a$  cv ct v f and + + + + + and each of the single recombinant classes, though neglecting recombinants between  $y^2$  and  $w^a$ , *i.e.* on recombination classes  $y^2w^a$  + + + +, + + cv ct v f,  $y^2w^a$  cv + + +, + + + ct v f,  $y^2w^a$  cv ct + +, + + + + v f,  $y^2w^a$  cv ct v + and + + + + + f. Counts were made on the offspring from each of the nine cultures and the mean numbers of coxal chaetae borne by the 10 males were used in the analysis as the basic data for that culture. Thus taking the three replicate cultures together, the chaeta counts of 30 males should be involved in each mean chaeta number for each parental or recombinant class from each of the three  $F_1$ 's. In

eight cases the number of males counted fell short of 30, being 28 in two cases, 27 in three cases and 25, 18 and 16 in one case each.

We used a process of differencing to arrive at an estimate of the effects on chaeta number of any gene differences between the tester chromosome, T, and the relevant line chromosomes, W, S or L, at loci linked to each of the marker genes. To take as an example the effects associated with the *cv* locus, the four recombinants  $y^2w^a + + + +$ ,  $y^2w^a cv + + +$ ,  $+ + + ct v f$  and  $+ + cv ct v f$  are the relevant classes. Denoting these four by a, b, c and d respectively, b differs from a only by having the gene *cv*, and any genes affecting chaeta number associated with it, from the T chromosome, while a has the + allele of *cv* and any chaeta genes associated with it from W, S or L as the case may be. Thus the difference in mean chaeta number  $a - b$  is an estimate of the effect of chaeta genes associated with the *cv* locus. Turning to c and d, d like b has *cv* and associated chaeta genes from T while c, like a, has them from W, S or L. Thus  $c - d$  also provides an estimate of the effects of chaeta genes associated with *cv* locus. Apart from sampling variation  $a - b$  differs from  $c - d$  only in respect of chaeta genes in other parts of the chromosome and these will contribute to the difference between the estimates if they interact with the chaeta genes associated with the *cv* locus in producing their effects. Since we have results from the three replicate cultures we can test for such interaction by finding the difference between the two estimates, *i.e.*  $a - b - c + d$ , from each replicate and test the significance of the departure from 0 of this difference summed over replicates against the inter-replicate variance. This can be done for all five marker genes (taking  $y^2$  and  $w^a$  as one since we have not used recombinants between them) for all three tests of W, S and L against T. No evidence of interaction has been found and it will therefore be neglected, though, of course, with so few d.f. the test is not very sensitive.

One further and important point must be made about this type of analysis. Consider, for example, a gene difference affecting chaeta number located between *cv* and *ct*. If the effect of the gene difference is  $d$  and its recombination value is  $p_1$  with *cv* and  $p_2$  with *ct*, we shall find an effect  $dp_2/(p_1 + p_2)$  on chaeta number associated with *cv* and  $dp_1/(p_1 + p_2)$  associated with *ct*, the divisor  $p_1 + p_2$  coming in because we are comparing the specified recombinant classes *cv* + in the one case and + *ct* in the other with the relevant non-recombinant classes *cv ct* and + +. We thus do not detect the full effect of the gene difference by its association with either *cv* or *ct*, the sum of the two effects being in fact the full effect of the gene difference. Where there are several chaeta genes located between the two marker loci, the effect associated with *cv* will be the sum of the  $dp_2/(p_1 + p_2)$  items, one from each chaeta gene, and correspondingly  $Sdp_1/(p_1 + p_2)$  will be the effect associated with *ct*.

Thus any one locus lying in this segment of the chromosome can give effects associated with both marker loci, any effect being the greater the closer the chaeta gene lies to the marker. So effects found associated with adjacent markers are not necessarily indicative of more than one gene affecting the character whose variation is under consideration. The results of an analysis such as we are undertaking must therefore be interpreted with caution. Indeed, this method of surveying a chromosome can only give a broad picture of the distribution of the activity along it. To obtain a more detailed picture and to locate the individual genes requires that the

survey be followed by an analysis of the more elaborate and more precise kind due to Thoday (1961).

As an example, the estimation of the difference in front coxal chaeta number between the S and T chromosomes associated with the cv locus is shown in table 3. The mean chaeta numbers of  $y^2w^a +++ +$  (a),  $y^2w^a cv +++ +$  (b),  $+++ ct v f$  (c) and  $+++ cv ct v f$  (d) are shown for the three replicate cultures with the number of males from which each mean is derived given in brackets after it. The three a-b differences are then found and their weighted mean is calculated. The weights reflect any differences in the number of males scored, that for a-b from replicate 1 for example being  $10 \times 6/10 + 6$ . The weighted mean for c-d is found similarly and combined with that of a-b by taking their simple average.

TABLE 3

*Activity in the cv region of the X chromosome: (a) assessing S-T, the difference between S and T; (b) W-T, S-T and L-T*

(a)		Mean chaeta number of class			
Replicate	a	b	c	d	
1	24.20 (10)	27.33 (6)	25.80 (10)	25.80 (10)	
2	25.14 (7)	26.40 (5)	24.00 (8)	27.30 (10)	
3	25.30 (10)	26.00 (7)	25.60 (10)	26.63 (8)	
Replicate	a-b	Weight	c-d	Weight	
1	-3.13	3.75	0.00	5.00	
2	-1.26	2.92	-3.30	4.44	
3	-0.70	4.12	-1.03	4.44	
Weighted mean difference	-1.70		-1.39		
(b)					
Chromosome	a-b	c-d	$\frac{1}{2}[(\overline{a-b}) + (\overline{c-d})]$	$\frac{1}{2}[(\overline{a-b}) - (\overline{c-d})]$	
W	-1.28	-0.63	-0.955	-0.325	
S	-1.70	-1.39	-1.545	-0.155	
L	-0.90	-1.20	-1.050	0.150	

This is entered in table 4 which sets out these average effects associated with the various marker genes for the differences between W and T, S and T, and L and T. We can of course compare the line chromosomes with one another through the intermediacy of T, since  $(W-T) - (S-T) = W-S$  etc.

Before taking the analysis further, however, we must return to the weighted means of a-b and c-d in table 3 and note that we combine them to give the overall average effect by finding  $\frac{1}{2}[(a-b) + (c-d)]$ . So  $\frac{1}{2}[(a-b) - (c-d)]$  will provide information about the error variance of this overall average. There are nine such values of  $\frac{1}{2}[(a-b) - (c-d)]$  for the cv region, one from each pair of legs from each of the three test, of W, S and L respectively against T. Assuming the error variances to be homogeneous over legs and lines, we thus have nine comparisons which can be combined to give an estimate of error variance based on 9 d.f. We can increase the number of comparisons on which the estimate is based by bringing in the information from regions of the chromosomes marked by other marker genes but we must not use adjacent markers as the information they yield is not independent in respect of the error variance any more than,

TABLE 4  
*Distribution of activity in the X chromosome: first experiment*

	$y_2^{2y^a}$			cy			ct			y			f		
	F	M	R	F	M	R	F	M	R	F	M	R	F	M	R
W-T	0.515	0.315	0.350	-0.955	-0.330	-0.660	-0.080	0.305	0.290	0.000	0.235	0.365	0.115	0.415	0.050
S-T	0.295	0.545	0.285	-1.545	-0.405	-1.055	0.220	0.500	0.390	0.570	-0.010	0.200	-0.630	0.300	0.300
L-T	4.180	2.185	0.850	-1.050	-0.465	-0.665	0.235	0.480	0.185	-0.330	-0.200	0.135	-1.000	0.465	-0.280
$\frac{1}{2}(W-S)$	0.110	-0.115	0.033	0.295	0.038	0.198	-0.150	-0.098	-0.050	-0.285	0.123	0.083	0.373	0.058	-0.125
$\frac{1}{2}(W+S-2T)$	0.405	0.430	0.318)**	-1.250***	-0.368	-0.858***	0.070	0.403	0.340	0.285	0.113	0.283	-0.258	0.358	0.175
$\frac{1}{2}(2L-W-S)$	3.775***	1.755***	0.533	0.200	-0.098	0.193	0.165	0.078	-0.155	-0.615	-0.313	-0.148	-0.743	0.108	-0.455

\* P = 0.05-0.01; \*\* P = 0.01-0.001; \*\*\* P < 0.001.

$\pm 0.3454$   
 $\pm 0.2442$   
 $\pm 0.4230$

as we have just seen, it is independent in respect of the effects of chaetae genes lying between them. We can use the data from *v* with that from *cv*, or that from  $y^2w^a$  with that from *ct*, but the data from *f* cannot be used as it is not independent of either of the other pairs. To be conservative we have taken the combination of the data from  $y^2w^a$  and *ct* as it is giving the higher estimate of error, which will of course be based on  $2 \times 9 = 18$  d.f. The number of comparisons has been increased further by bringing in similar data from the results in respect of  $y^2$  and *ec* in a second experiment which is described below. The figures from the two experiments showed no heterogeneity in respect of error variance and pooling them yields an estimate  $V = 0.119311$  for 36 d.f.

The results of the first experiment, with which we are now concerned, are set out in the upper part of table 4 in the form that they were obtained, *viz.* as the differences between the *W*, *S* and *L* chromosomes and *T*. We can, however, replace these original differences  $W-T$ ,  $S-T$  and  $L-T$  by three further comparisons among them which present the information they contain in a more useful way. Various sets of three comparisons could be chosen for this purpose but the set we have chosen is

$$(i) \frac{1}{2}[(W-T) - (S-T)] = \frac{1}{2}(W-S),$$

comparing the *W* and *S* chromosomes,

$$(ii) \frac{1}{2}[(W-T) + (S-T)] = \frac{1}{2}(W+S-2T)$$

comparing the average effect of *W* and *S* with *T*, and

$$(iii) \frac{1}{2}[2(L-T) - (W-T) - (S-T)] = \frac{1}{2}(2L-W-S)$$

comparing *L* with the average of *W* and *S*. Comparison (i) is statistically independent of (ii) and (iii), but (ii) and (iii) are not fully independent of one another. We have, however, used them, for despite their statistical shortcomings they yield the genetical estimates that we want.

The values for these three comparisons are given in the lower part of table 4 together with their standard errors. We have seen that the original differences,  $W-T$  etc., are subject to an error variance  $V = 0.119311$ , giving a standard error  $\sqrt{0.119311} = \pm 0.3454$ . The error variance of  $\frac{1}{2}(W-S) = \frac{1}{2}[(W-T) - (S-T)]$  and  $\frac{1}{2}(W+S-2T) = \frac{1}{2}[(W-T) + (S-T)]$  will clearly both be  $\frac{1}{4}.2V = 0.059656$  and their standard error  $\sqrt{0.059656} = \pm 0.2442$ . The error variance of

$$\frac{1}{2}(2L-W-S) = \frac{1}{2}[2(L-T) - (W-T) - (S-T)]$$

will however be  $\frac{1}{4}.6V = 0.178967$  and its standard error 0.4230.

The first thing to observe from the table is that none of the 15 entries for  $\frac{1}{2}(W-S)$ , one from each pair of legs in respect of each marker locus, is significant. There is thus no evidence of any genetic difference between *W* and *S* in respect of coxal chaeta numbers. They may therefore, be combined for comparisons with *T* and *L*, which is the reason for the choice of the set of comparisons we have used. Turning to the comparison of *W* and *S* jointly with *T*, we find that the numbers of chaetae on the *F* and *R* legs are significantly lower with *W* and *S* than with *T* in respect of the *cv* region of the chromosome. The *W* and *S* value for the *M* legs is not significantly lower than with *T* but it is deviating in the same direction as for *F* and *R*, its

probability is 0.20 and if we take F, M and R values together they are jointly significant. There can thus be little doubt that W and S differ from T at one or more loci near to *cv*, with alleles mediating a lower chaeta number than T. No other entry in the  $\frac{1}{2}(W+S-2T)$  line is of itself significant, but we may note that in respect of the  $y^2w^a$  region all three legs show high positive values that for M being nearly significant with  $P = 0.08$  while even the lowest of them for R, has a probability 0.20. Furthermore, when taken together the sum of the values from F, M and R is highly significant with  $P = 0.01$ . Again there can be little doubt that W and S differ from T at the left end of the X chromosome though here in contrast to the *cv* region they carry allele(s) mediating a higher number of chaetae than T. The only other region worthy of note is that marked by *ct*, where M and R show fairly large (although falling just short of significance) values and the sum of the differences on F, M and R just fails to achieve formal significance with  $P$  just over 0.05. There is thus a hint that W and S differ from T in the region *ct* also, and if the difference is genuine it cannot be just a reflection of the effect detected by *cv* since the two act in opposite directions: they must be due to different genes.

Turning finally to the difference between L on the one hand and W and S on the other, we observe that apart from the  $y^2w^a$  region there is no real indication that  $\frac{1}{2}(2L-W-S)$  departs from 0 except perhaps for the F legs in the *v* and *f* regions, the value in respect of *v* having  $P = 0.15$  and that in respect of *f* having  $P = 0.08$ . The values are in the same direction and could be ascribed to a gene or genes lying between *v* and *f* with the allele(s) in L giving a lower chaeta number than those in W and S. We shall have occasion to refer again to this possible difference in the *v* and *f* region.

There is no ambiguity about the results from the  $y^2w^a$  region.  $\frac{1}{2}(2L-W-S)$  is highly significant for the F and M legs and while the value for R is not significant the sum of F, M and R is also highly significant with  $P$  very small indeed. Since in this region L thus exceeds W and S in its genic capacity for coxal chaeta production and W and S similarly exceed T, the comparisons among the four chromosomes indicate at least a double difference (*i.e.* either three alleles at one locus or differences at two loci) in respect of chaeta production.

The left end of the X was investigated further in a second experiment in which an additional marker gene *ec* was introduced into T, which thus carried  $y^2w^a ec cv ct v f$ . Both this chromosome and the  $y^2w^a cv ct v f$  chromosome of the first experiment were made up from two chromosomes marked respectively by  $y^2w^a cv v f$  and  $y ec ct v f$ . Though the two T chromosomes we have used obviously differ at the *ec* locus, and may do so at loci lying close to *ec*, the method by which they were constructed makes it very likely that they were otherwise closely similar if not identical. Methodologically the second experiment was carried out in just the same way as the first, and its results were also analysed in the same way. The difference between them was that whereas the first experiment was aimed at providing a general survey of the X chromosome in relation to coxal chaeta production, the second was aimed at a closer analysis of the left end of the chromosome. The single recombinants between  $y^2$  and  $w^a$ ,  $w^a$  and *ec*, and *ec* and *cv* were thus picked out for observations of their chaeta numbers; but recombinants to the right of *cv* were discarded, the whole chromosome to the right of *ec* thus being treated as a unit marked simultaneously by *cv*,

ct, v and f. The maximum number of males counted for a genotype was 60, that is twice the number in the first experiment. Because, however, the loci of  $y^2$ ,  $w^a$  and  $ec$  are close together, recombination between them (especially between  $y^2$  and  $w^a$ ) is relatively rare. In consequence the number of males recovered did not always reach 60 and in the  $+w^aec\ cv\ ct\ v\ f$  class was always rather low, being 17 in each of the S and L analyses and no more than nine from the W cross. This must of course reduce the precision of the estimates of the mean numbers of chaetae in these classes; but it will be recalled that the estimate of error variance is not based directly on the means themselves, but on the differences between the  $a-b$  and  $c-d$  estimates of the chaeta differences associated with the individual markers. As we have already noted, these estimates of error variance were found to be homogeneous over the two experiments and so were pooled to provide a common estimate of error for use with all the markers in both experiments.

The results from the second experiment are set out in table 5, with levels of significance indicated in the same way as in table 4. Just as in the first experiment there is no evidence of differences in chaeta number between W and S in respect of any of the marked regions. The  $\frac{1}{2}(2L-W-S)$  comparison, however, shows a significant excess of L over W and S associated both with  $y^2$ , in respect of the F legs, and with  $w^a$  in respect of all three pairs of legs, just as there was with  $y^2w^a$  in respect of F and M in the first experiment. Furthermore, there is some indication of the same kind of difference associated with  $ec$ —perhaps because a gene contributing to its production lies between  $w^a$  and  $ec$ . We should note, however, that even if this is the correct interpretation, it could not account simultaneously for the effect associated with  $y^2$ , which must therefore require a further gene between  $y^2$  and  $w^a$ . At the same time the excess of W and S over T, shown by  $\frac{1}{2}(W+S-2T)$  associated with the  $y^2w^a$  region in the first experiment is displayed by neither  $y^2$  nor  $w^a$  in the second; but such an excess is shown by  $ec$ .

The length of chromosome effectively marked by  $y^2w^a$  and  $ec$  must be much the same as that earlier marked by  $y^2w^a$  since  $ec$  is much closer to  $w^a$  than to  $cv$ . Thus, apart from any genes newly brought into T along with  $ec$ , the sum of the corresponding effects associated individually with  $y^2$ ,  $w^a$  and  $ec$  in the second experiment should approximately equal the effects associated with  $y^2w^a$  jointly in the first. The sums of the  $y^2$ ,  $w^a$  and  $ec$  effects are shown in the lower part of table 5. Being the sum of three observations they will have an error variance three times that of the single items and standard errors  $\sqrt{3}$  times as large. Their summed effects will thus be less precise than the single differences associated with  $y^2w^a$  in the first experiment (table 4). When the appropriate tests of significance are applied we can see that the summed  $y^2$ ,  $w^a$  and  $ec$  effects from the second experiment are fully consistent with the  $y^2w^a$  effects from the first; in no case do corresponding effects differ significantly between the experiments. Furthermore, the pattern the  $\frac{1}{2}(2L-W-S)$  difference shows over the three pairs of legs is the same in both experiments, the figures for F and M showing in fact an almost surprising consistency and even those for R, though not agreeing so closely, not only do not disagree significantly but show a difference less than its standard error. The pattern of  $\frac{1}{2}(W+S-2T)$  is not so strikingly similar in the two experiments, it is higher on M and R in the second than the first but lower on F. Even so their agreement is fully satisfactory in that

TABLE 5  
Distribution of activity in the X chromosome: second experiment

	$y^2$			$w^d$			cc			Rest		
	F	M	R	F	M	R	F	M	R	F	M	R
W-T	-0.354	0.369	-0.159	0.551	-0.028	0.143	0.096	0.768	0.965	-0.542	0.725	-0.192
S-T	-0.176	0.154	0.477	-0.116	-0.085	-0.533	-0.119	0.412	0.773	-1.412	0.339	-0.230
L-T	0.748	0.163	-0.133	2.219	1.207	1.029	0.834	1.366	1.250	-2.774	-0.075	-1.300
$\frac{1}{2}(W-S)$	-0.089	0.108	-0.318	0.334	0.029	0.338	0.108	0.178	0.096	0.435	0.193	-0.081
$\frac{1}{2}(W+S-2T)$	-0.265	0.262	0.159	0.218	-0.057	-0.195	-0.012	0.590*	0.869***	-0.977***	0.532*	-0.311
$\frac{1}{2}(2L-W-S)$	1.013*	-0.099	-0.292	2.002***	1.264**	1.224**	[0.846	0.776	0.381]**	[-1.797***	-0.607	-0.989*]**
	$y^2+w^d+cc$											
	F	M	R									
W-T	0.293	1.109	0.949	} $\pm 0.5983$								
S-T	-0.411	0.481	0.717									
L-T	3.801	2.736	2.146									
$\frac{1}{2}(W-S)$	0.352	0.314	0.116	} $\pm 0.4230$								
$\frac{1}{2}(W+S-2T)$	-0.059	0.795	0.833									
$\frac{1}{2}(2L-W-S)$	[3.860***	1.941**	1.313]**	} $\pm 0.7327$								

\* P = 0.05-0.01; \*\* P = 0.01-0.001; \*\*\* P < 0.001.



the corresponding items agree well within the limits of error variation and also in that the  $\frac{1}{2}(W+S-2T)$  effects in both experiments, though in the same direction as  $\frac{1}{2}(2L-W-S)$ , are less than them. Evidently the second experiment is picking up the same differences as the first, but also providing the further information that at least two loci are involved, one between  $y^2$  and  $w^a$  and the other either between  $w^a$  and  $ec$  or just to the right of  $ec$ , rather than only one locus with three alleles as was also a possible interpretation of the results from the first experiment. Again we note that the difference between  $W$  and  $S$  and  $T$  earlier associated with  $y^2w^a$  but picked up by  $ec$  in the second experiment, cannot be a reflection of that associated with  $cv$  since they are in opposite directions.

The behaviour of the rest of the chromosome, from  $cv$  to  $f$ , in the second experiment requires little comment. It confirms the difference between  $W$  and  $S$  on the one hand and  $T$  on the other which was earlier detected as associated with  $cv$ . It also reveals a significantly negative value for the comparison  $\frac{1}{2}(2L-W-S)$  on both  $F$  and  $R$  legs which did not emerge in the first experiment. We should note, however, that both the  $v$  and  $f$  regions gave such a negative value for  $F$  and  $R$  in the earlier experiment. We must therefore be detecting the same effect, significance appearing in the second experiment because the two regions are taken jointly. There is thus evidence of a locus (or loci) between  $v$  and  $f$  which is associated with a lower chaeta number in  $L$  than in  $W$  and  $S$ . Though not significant, the difference for the chaetae on the  $M$  legs is in the same direction and we must assume it to be reflecting the action of the same gene(s) that produces the  $F$  and  $R$  differences.

The results of the two experiments are thus in good agreement. They reveal that at least four and possibly six or seven loci were detectable in the  $X$  chromosome as affecting coxal chaeta number, with something of a concentration of them at the left end of the chromosome. As we have already noted Thoday's more searching methods would be needed to take the analysis further.

### 3. VARIATION IN GENE ACTION

The experiments described in Section 2 show that all three major chromosomes carry genes affecting the numbers of coxal chaetae on the three pairs of legs, and that such genes are to be found at a number of loci in the  $X$  chromosome. These findings arise from comparisons among the inbred lines and the genes so detected are thus genes in respect of which differences are to be found among the various lines. Comparisons among the three pairs of legs on the same fly or on flies of the same genotype, provide a different type of evidence which will reflect the effects not only of genes in respect of which differences occur between the lines, but also those in which the lines are all alike.

Consider a pair of inbred lines differing in a gene pair  $A-a$ , one line being  $AA$  and the other  $aa$ . Let  $AA$  and  $aa$  be associated with the production of  $A$  and  $a$  chaetae respectively on a pair of legs. Then given a uniform genetic background, the mean chaeta number of the two lines will be  $\frac{1}{2}(A+a)$  and the deviations of the two lines from this mean will be  $\frac{1}{2}(A-a)$  and  $-\frac{1}{2}(A-a)$  respectively. In the notation of biometrical genetics  $\frac{1}{2}(A+a)$  is a component of  $m$ , the mid-parent value, and  $\frac{1}{2}(A-a)$  is  $d_a$  measuring

the difference in effect of the two alleles. If the two do not differ, at least in respect of their effects on this character,  $A = a$  and  $d_a = 0$ , but the locus still contributes  $A$  to  $m$  the mean expression of the character. Thus  $m$  is affected by genes which do not differ among the lines, though  $[d] = S(d_a)$  is not affected by this.

Now if  $m$  differs from leg to leg, as in Section I we saw it to do, the genes must be changing in the effects of their action with the changes in the situation in which they find themselves as a result of the processes of differentiation during the development of the fly, and in particular of its legs—processes which must of course themselves be under genetic control. If  $m$  is affected we must thus in general expect  $[d]$  to be altered, though what the relationship of the changes in  $m$  and  $[d]$  will be must depend on how the changes in situation from leg to leg affect the effective action of the genes. Where all the genes are affected alike in their expression, as might for example be expected with alteration in the amount of a raw material necessary for them to express themselves, we should expect  $[d]$  to alter proportionately with  $m$ . If, however, the genes are affected differentially by the changes in situation, as might be the case where the action of some is reduced or virtually eliminated while others are relatively unaffected, we should not expect such a simple relationship between  $m$  and  $[d]$ . To put it in other words if the rise in average chaeta number from the R through the M to the F legs is due to the set of genes showing an increase in their effective action while retaining the same balance of effect among themselves,  $[d]$  should rise proportionately to  $m$ ; but if in effect, because of the operation of some controlling system, the rise in  $m$  is due to an increase in the number of genes acting effectively,  $[d]$  will not show a change proportionate to  $m$ .

We have two pieces of evidence that the changes between the R, M and F legs do not involve all the relevant genes equally. In the first place, if they did the position would be formally similar to a genotype  $\times$  environment interaction of the multiplicative type analysed by Mather (1975). In such a case where we have a number of genotypes placed in a series of different environments or situations, and we plot the expression of the character for each individual genotype against the expression averaged over all genotypes, environment by environment, we obtain a pencil of straight lines of differing slopes but all passing through the origin. This is easily tested by fitting a straight regression line to each genotype with the restriction that it must pass through the origin, and testing the residual variation of the R, M and F points for the genotypes round their regression lines, using an appropriate error variance obtained from replicate observations. This has been done separately for the three sets of substitution lines, W/S, S/L and L/W and also for the set of 18 Texas inbred lines. The results are set out in table 6. In every case the residual variation round the straight regression lines is highly significant, as shown by the  $\chi^2$  test. The regression cannot in fact be regarded as straight and  $[d]$  cannot therefore be rising proportionately with  $m$ . The chief departures spring from the F legs, but even when these are excluded the observations on R and M in most cases still do not agree with straight regressions passing through the origin, and the rise in  $[d]$  from R to M cannot be accepted as proportional to the rise in  $m$ .

The second piece of evidence comes from the substitution lines and is of a different kind. The analyses of variance summarised in table 3 suggest that the balance of the contributions from chromosomes X, II and III to the

TABLE 6  
*Test of equal enhancement of gene effects*

		18°C	25°C
Texas lines	$\chi^2_{[34]}$	157.60	165.50
	P	v.s.	v.s.
W/S	$\chi^2_{[14]}$	48.27	28.45
	P	v.s.	0.02-0.01
S/L	$\chi^2_{[14]}$	147.86	202.43
	P	v.s.	v.s.
L/W	$\chi^2_{[14]}$	84.55	90.89
	P	v.s.	v.s.

variation among the eight substitution lines of a set varies from leg to leg. Thus in the W/S set of lines the variation of the F legs is attributable almost entirely to III, but that of the R legs almost entirely to II. In the same way, among the lines of the L/W set II makes a larger contribution than either X or III to the variation shown by the R legs, but almost nothing to that of the M legs, to which both X and III contribute significantly, while chromosomes II and III contribute almost equally to the variation on F though both materially less than X. Such a changing balance from leg to leg of the contributions to the variation is not compatible with  $[d]$  changing proportionately to  $m$ . We can, however, take the analysis further by finding the differences ascribable to the different chromosomes in the three sets of lines.

Now the excess in mean number of chaetae ascribable to the X chromosome of the W line over the mean of all the W/S set of lines can be found as:  $\frac{1}{8}[\overline{WWW} + \overline{WWS} + \overline{WSW} + \overline{WSS} - \overline{SWW} - \overline{SWS} - \overline{SSW} - \overline{SSS}]$  where  $\overline{WWW}$  is the mean chaeta number of the lines whose constitution is WWW. This excess will reflect the action of all the genes carried by the X chromosome and so is strictly a  $[d]$ . We will, however, designate it as  $d_{WSX}$  for convenience. The departure of the X from S over the mean is similarly  $d_{SWX} = -d_{WSX}$ . In the same way  $d_{WS2} = -d_{SW2}$  measures the departure ascribable to chromosome II and so on. When we move to the S/L set of lines we can similarly find  $d_{SLX} = -d_{LSX}$ ,  $d_{SL2} = -d_{LS2}$  etc. and the L/W set of lines yield  $d_{LWX} = -d_{WLX}$  etc. These  $d$  values are of course directly related to the corresponding mean squares of table 2. They are set out in table 7.

While  $d_{WSX}$  provides direct information about the difference in chaeta producing power of the X from W over that from S, further information can be obtained about it indirectly from the other two substitution sets. These compare W with L and S with L, and we can thus find

$$d_{WSX} = d_{WLX} - d_{SLX} = -(d_{SLX} + d_{LWX}).$$

Being based on the sum of two  $d$  values this second estimate will have twice the variance and so yields half the information of the direct estimate  $d_{WSX}$ . Our best estimate of the difference in chaeta producing power between the X's from W and S is thus

$\hat{d}_{WSX} = \frac{1}{3}(2d_{WSX} - d_{SLX} - d_{LWX}) = \frac{1}{3}[2(-0.541) + 1.331 - 1.708] = -0.486$  at 25°C from the data of table 7. The values of  $\hat{d}_{WS}$ ,  $\hat{d}_{SL}$  and  $\hat{d}_{LW}$  for all three chromosomes at both temperatures are given in the lower part of the table.

As we have seen  $d_{WSX}$  is a measure  $\bar{X}_W - \bar{X}_S$ ,  $d_{SLX}$  a measure of  $\bar{X}_S - \bar{X}_L$  and

TABLE 7  
*Differences in activity of W, S and L, by chromosomes and legs at 25°C and 18°C*

	X			II			III			
	F	M	R	F	M	R	F	M	R	
25°C	$d_{WS}$	-0.541	-0.021	-0.147	-0.193	0.185	-1.153	-0.674	-0.053	
	$d_{SL}$	-1.331	-0.688	0.056	0.100	-0.159	2.125	-0.100	0.416	
	$d_{LW}$	1.708	0.967	0.580	0.173	-0.555	-0.961	0.733	0.483	
18°C	$d_{WS}$	-0.047	0.463	0.197	0.250	0.654	-0.874	0.038	0.091	
	$d_{SL}$	-1.209	-0.966	-0.265	-0.341	-0.163	1.409	-0.616	-0.381	
	$d_{LW}$	1.927	1.023	1.020	0.152	-0.338	-0.748	0.561	0.056	
				All subject to $s_d = \pm 0.2141$						
25°C	$d_{WS}$	-0.486*	-0.163	-0.310	-0.220	0.361	-1.157***	-0.660**	-0.058	
	$d_{SL}$	-1.276***	-0.746***	-0.107	0.073	0.017	2.121***	-0.086	-0.421*	
	$d_{LW}$	1.763***	0.909***	0.417*	0.146	-0.379*	-0.965***	0.747***	0.478*	
18°C	$d_{WS}$	-0.271	0.290	-0.120	0.230	0.603**	-0.803***	0.044	0.169	
	$d_{SL}$	-1.433***	-1.139***	-0.582**	-0.361	-0.214	1.480***	-0.610**	-0.303	
	$d_{LW}$	1.703***	0.850***	0.703***	0.132	-0.389*	-0.677**	0.567***	0.134	
				All subject to $s_d = \pm 0.1748$						

\* P = 0.05-0.01; \*\* P = 0.01-0.001; \*\*\* P = <0.001.

Underlines indicate significant change in sign of  $d$  between legs.

$d_{LWX}$  a measure of  $\bar{X}_L - \bar{X}_W$ . Thus  $d_{WSX} + d_{SLX} + d_{LWX} = 0$  apart from sampling error, and its departure from 0 provides a measure of sampling error. We can obtain 18 such values, one from each chromosome, from each leg at each temperature. Summing the squares of the 18 observed deviations from 0, using the  $d$  values of table 7, gives 2.4754 as the sum of squares for 18 d.f. from which we obtain 0.13752 as the mean square estimating the error variance of  $d_{WS} + d_{SL} + d_{LW}$ . Since this is the error variance of the sum of the three  $d$  values, the error variance of a single  $d$  must be

$$V_d = 0.13752 \div 3 = 0.045841.$$

Now since  $\hat{d}_{WS} = \frac{1}{3}(2d_{WS} - d_{SL} - d_{LW})$  and similarly for  $\hat{d}_{SL}$  and  $\hat{d}_{LW}$  we can find  $V_d = \frac{1}{6}(4 + 1 + 1)V_d = \frac{2}{3}V_d = 0.03056$ , giving  $s = 0.1748$  as noted in the table. The significance of the  $\hat{d}$  values can now be tested by finding  $t = \hat{d}/s_d = \hat{d}/0.1748$ . The level of significance of each  $\hat{d}$  is indicated in the table.

So far the analysis is effectively little more than an alternative form of the tests of significance of the analyses of variance in table 2, using an empirical estimate of error variance and treating data from the two temperatures separately. We can, however, take a further step crucial to our present consideration, since in the present analysis the signs of the  $d$ 's are available. If we look at  $\hat{d}_{LW}$  for chromosome III at 25°C we see that it is significant for all three legs, indeed having a probability below 0.001 for M and F. But whereas it is significantly positive on R and M, it is significantly negative for F. In other words the chromosome III from L mediates a higher chaeta number than its homologue from W on the R and M legs, but a lower one on F, even though of course  $m$  has risen from R to M to F. The same is true of  $\hat{d}_{LW}$  at 18°C, thus providing independent confirmation. Basically the same picture is shown by chromosome II of L and W, albeit at a lower level of significance, but here, however, the change of sign in  $\hat{d}_{LW_2}$  is more likely to be between R and M than between M and F as it was with  $\hat{d}_{LW_3}$ . There is also a significant change of sign in  $\hat{d}_{SL}$  for chromosome III. This, however, is no more than a reflection of the change we have already noted in  $\hat{d}_{LW_3}$ . Owing to the structure of the experiment we expect a change revealed by one comparison, say  $\hat{d}_{LW}$ , to be mirrored by a change in another comparison, say  $\hat{d}_{SL}$ , the change being in opposite directions in the two cases, as indeed we observe to be the case. Thus the comparisons for a given chromosome at a single temperature are not independent; but those for the different chromosomes at the same temperature or for the same chromosome at different temperatures are independent. And we find that the results from the two temperatures confirm one another in showing that there are changes of sign in two of the three chromosomes.

Such changes of sign completely rule out the possibility that  $[d]$  is changing proportionately with  $m$ , and with it any possibility that the genes are changing similarly to one another in their effective activities, as we proceed from the R through the M to the F legs. Rather they suggest that genes which show very little if any activity on one pair of legs are fully active on another. And since the chaeta number rises from R to M and then from M to F it suggests that more genes are effectively active in respect of the M legs than the R, and more again in respect of the F legs than the M. This could account fully for the results from the substitution lines as is illustrated by the hypothetical example in fig. 3. The  $\alpha$  genes of the figure

are active on all three pairs of legs, being associated in one line (J) with the production of 16.5 chaetae on average and in the other line (K) with the production of 15.5 chaetae. This gives  $\frac{1}{2}(16.5+15.5) = 16.0$  chaetae as their contribution to  $m$  and  $\frac{1}{2}(16.5-15.5) = 0.5$  as their  $[d]$ , line J carrying the alleles which on balance give the higher chaeta number. The  $\beta$  genes are not active on R but are active on both M and F. They contribute

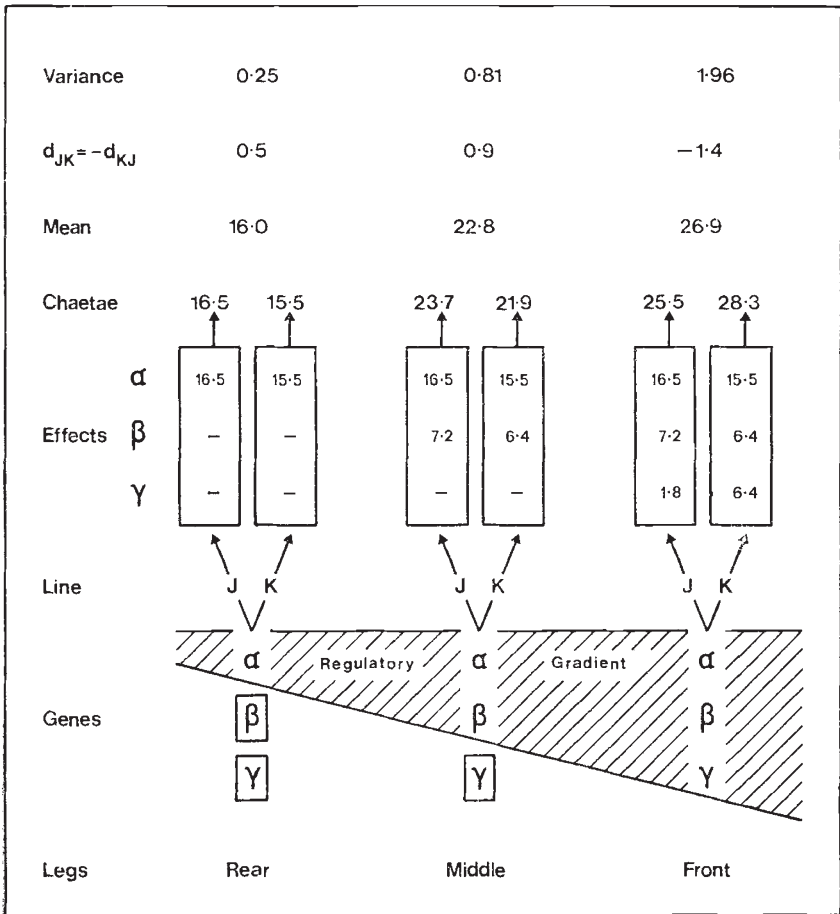


FIG. 3.—The postulated mode of determination of the gradient in number of coxal chaetae. The genes fall into three classes,  $\alpha$  genes which produce their effects on all three pairs of legs,  $\beta$  genes which are ineffective on R but effective on M and F legs, and  $\gamma$  genes which are effective on the F legs only. The genes in each set differ between the two lines J and K. The average number of chaetae borne on any pair of legs is determined by the summed effects of such sets of genes as are effective at that level. The numbers of chaetae for each pair of legs in each line are shown, together with the mean numbers, the values of  $d_{JK}$  and the variances derived from the  $m$ . The value of  $d_{JK}$  changes sign between the M and F legs because J carries the more powerful alleles of sets  $\beta$  and  $\gamma$ , but K carries the more powerful alleles of set  $\alpha$ . Note that the genes of a set are not all carried on the same chromosome but can be distributed over all the chromosomes. The figure also shows how the differences in effect of the three sets of genes could arise from differences in their patterns of activity reflecting their differential responses to a gradient of regulatory substance.

$\frac{1}{2}(7.2+6.4) = 6.8$  to  $m$  which thus rises from 16.0 to 22.8 on the M legs. They contribute  $\frac{1}{2}(7.2-6.4) = 0.4$  to  $[d]$ . Again the J line has the higher alleles on balance and  $[d]$  rises from 0.5 to  $0.5+0.4 = 0.9$ . We may note in passing that if none of the  $\beta$  loci had different alleles in the J and K lines, contributing in fact 6.8 chaetae in both lines,  $m$  would still have risen to 22.8 but  $[d]$  would have remained at 0.5. Finally there are the  $\gamma$  genes, inactive on both R and M legs, but active on F. Here, in the example, the K line carries on balance higher alleles than J, its  $\gamma$  genes mediating the production of 6.4 chaetae while those of J mediate the production of only 1.8. They will thus contribute  $\frac{1}{2}(1.8+6.4) = 4.1$  to  $m$ , which rises from 22.8 to  $22.8+4.1 = 26.9$ ; but their contribution to  $[d]$  is  $\frac{1}{2}(1.8-6.4) = -2.3$  and  $[d]$  thus falls to  $0.9-2.3 = -1.4$  and so changes sign in passing from M to F. The resemblance of the changes observed in  $d_{LW}$  for chromosome III to those of  $[d]$  in the example requires no emphasis: clearly the hypothesis can account for the observations. Finally we might also note that if the  $\gamma$  genes of the illustration had resembled the  $\alpha$  and  $\beta$  genes in having their higher alleles in J rather than K, they would have contributed 2.3 to  $[d]$  which would thus not have fallen, but risen from 0.9 to 3.2 in passing from M to F. The picture would thus have resembled the results for  $d_{LW}$  in respect of the X chromosome in showing a steady rise of  $[d]$  accompanying the rise in  $m$  from rear to front.

#### 4. DISCUSSION

Differential gene activity (or inactivity if we choose to read the system from front to rear rather than rear to front) will account for all the observations that we have made, and indeed is required by many of them. Thus it not only provides an interpretation of the situation where  $[d]$  rises *pari passu* with  $m$  as we proceed from rear to front, which might however also be interpreted in terms of a general enhancement of gene effects, but it also provides an interpretation of relations between  $[d]$  and  $m$  which cannot be accounted for in terms of simple enhancement. We have already seen how differential activity can account for a change in the sign of  $[d]$  between legs, as has been observed with chromosome II and III in the chromosome assays from the substitution lines. Other patterns are obviously possible and some are indeed suggested, even if not always statistically settled, by the X chromosome analysis. A whole chromosome might be expected to carry a number of genes affecting the character in question and its overall effects will be the resultant of all the genes concerned. But if we look at the effects of segments of chromosome, as we are doing in the X chromosome analysis, we might be getting at least nearer to seeing the behaviour of single gene differences (or at any rate the single gene differences whose effects stand out by their greater size from those of their fellow genes) especially where we are observing small segments as at the left end of the X in the analysis. In such cases we might hope, within the limits set by error variation, to recognise the type of gene  $\alpha$ ,  $\beta$  or  $\gamma$ , whose effects we are observing. Thus the genes revealed by the  $\frac{1}{2}(2L-W-S)$  comparison in the  $y^2$  and  $w^a$  regions in table 5 suggest a near uniform difference over the R and M legs but a much increased one on the F legs. This would imply the action of a  $\gamma$  gene to account for the change from M to F, superimposed on an  $\alpha$  gene producing the value common to R and M. The same would be the case with the

difference in the v-f region revealed by the first analysis and perhaps that shown by the  $\frac{1}{2}(W+S-2T)$  comparison in the cv region (table 4). A significant value of a comparison that was uniform over all three pairs of legs, such as is suggested by  $\frac{1}{2}(W+S-2T)$  in the  $y^2w^a$  region of the first experiment (though not confirmed by the second experiment) would imply the action of an  $\alpha$  gene.

These results from the X chromosome analysis are suggestive though nothing more. A more confident classification of the types of genes involved in the differences between the W, S, L and Tester X chromosomes would require larger and more precise experiments. Other predictions are, however, possible from our hypothesis of differential gene activity. First, since variation on the R legs arises solely from  $\alpha$  gene differences, that on the M legs from the same  $\alpha$  genes together with  $\beta$  genes, and that on the F legs from the same  $\alpha$  and  $\beta$  genes to which  $\gamma$  gene variation is now added, one would expect the variation among genotypes in respect of coxal chaeta number to rise from rear to front. Table 8 shows, leg by leg, the numbers of coxal chaetae averaged over the 18 inbred lines stemming from the Texas population, together with the variance of the lines round their common mean. As expected the variance is lowest on the R, intermediate on the M and highest on the F legs. This might, of course, be a scalar effect, but it is not one which a coefficient of variation would remove.

TABLE 8  
*Mean and variance of chaeta numbers of the 18  
Texas lines, by legs*

	F	M	R
Mean	23.2	19.9	15.1
Variance	6.28	2.23	0.73

The second kind of prediction relates to the effects of selection. A number of such predictions can be made of which we now need to note only one. If, for example, we select for a change in the difference between the chaeta numbers of F and M legs, any response which comes about must be due to alteration of the chaeta number of the F legs, the chaeta number of the M legs remaining essentially as they were. This is because any alteration of chaeta number of M must be due to change in the  $\alpha$  and  $\beta$  genes which also equally effect the chaeta number on the F legs. A rise or fall in the F-M difference must thus be due to  $\gamma$  genes whose effects are displayed only on F. We have carried out such selection experiments, but an account of them and their results must be reserved for a later publication.

While the results, particularly the change in sign between legs of the [*d*] displayed by a pair of homologous chromosomes, show that different genes must have different patterns of activity on the various legs, we cannot demonstrate from these observations that a gene shows either no activity or full activity according to its type and the leg in question. The difference between the levels of activity on the different legs must be substantial or we would hardly expect to observe changes as drastic as a reversal in the sign of [*d*]. At the same time there is nothing to show that it is as absolute as we have assumed. This clearly is a matter on which further information is needed: we have adopted the assumption that the activity of a gene is all



or none because this is the simplest postulate and hence the one most easily tested. And even should the distinction turn out to be less than fully absolute our basic hypothesis would still stand, because its essential requirement is the differential changes in the activity of the genes in that some of them will be relatively inactive on legs where others are much more active, while on other legs all of them will show high activity.

It is tempting to relate the classes of genes we have inferred to the differences in distribution of the chaetae themselves on the coxae of the three kinds of legs (fig. 1). On the R legs the chaetae are borne round the periphery of the coxa, but on the M and even more on the F legs these peripheral chaetae are supplemented by others arising nearer to the centre of the coxa. This suggests that  $\alpha$  genes may be concerned with the production of peripheral chaetae, and the  $\beta$  and  $\gamma$  genes with those more centrally placed. We have endeavoured to carry out separate genetical analyses of peripheral and central chaeta numbers, but these have failed to produce any meaningful partition of the differences from which the classes of gene have been inferred. At the same time, this failure could have arisen from the difficulty of distinguishing unambiguously between the two kinds of chaeta, the distinction between which is less easy than fig. 1 might suggest.

However this may be, the question remains of how the differential activity is determined. We can only speculate as to the answer. It may be that all the genes are active on all legs, but that their products are not all equally effective in the production of chaetae under the different circumstances prevailing on the different legs. This could arise if the three classes of gene required different raw materials with which their products could react, and these raw materials themselves showed different patterns between the legs in their availability, whether because there was less produced in the cells of some legs than others or because of differences in the incidence of competitive reactions that also depended on them, but expressed their utilisation in ways other than chaeta production. In either case we are left postulating further genes concerned with the production of the materials or their competitive utilisation, and asking how these further genes can come to act differentially between the legs.

A different and more attractive view is that the differential activity of the chaeta genes springs from differential activation, the  $\alpha$  genes being themselves activated in all the legs, the  $\beta$  genes not activated in the R but activated in M and F, and the  $\gamma$  genes silent in both R and M but activated in F. This would suggest that the structural genes responsible for the mediation of chaeta production are subject to control, in the sense of being subject to activation or non-activation by further elements which in turn respond to a regulatory material or materials whose presence or concentration varied between the legs. It might be noted that controlling elements of the kind thus postulated have been reported in relation to the "rosy" locus in *Drosophila* (Chovnick *et al.*, 1976). The regulatory material would also presumably be itself gene controlled in its production and distribution, to give the gradient from front to rear that would be sufficient to account for our results. A tripartite system of this kind closely parallels the system which has been proposed and discussed by Britten and Davidson (1969) and Davidson and Britten (1973).

The changes in coxal chaeta number brought about by the reduction of temperature from 25°C to 18°C illustrated in fig. 2, are simply accounted

for by a change in a regulatory gradient. It is, however, impossible to obtain direct evidence of a gene controlled regulation from data such as our observations have yielded. Our only indication of it was the occurrence of what must have been a mutant in one of the inbred lines which presumably altered the shape of the gradient, since in addition to lesser changes in the chaeta numbers of the R and M legs it so lowered the chaeta number of the F legs as to reduce their average below that of M. Unfortunately, the mutant was lost before its type of inheritance could be established or its effects on chaeta number worked out in full detail.

*Acknowledgments.*—We gratefully acknowledge financial support from the Leverhulme Trust Fund and the Agricultural Research Council, and also valuable comments and suggestions from our colleague Dr P. D. S. Caligari.

## 5. REFERENCES

- BRITTEN, R. J., AND DAVIDSON, E. H. 1969. Gene regulation for higher cells: a theory. *Science*, *165*, 349-357.
- CALIGARI, P. D. S., AND MATHER, K. 1975. Genotype-environment interaction III. Interactions in *Drosophila melanogaster*. *Proc. Roy. Soc. B*, *191*, 387-411.
- CHOVNICK, A., GELBART, W., MCCARRON, M., OSMOND, B., CANDIDO, E. P. M., AND BALLIE, D. L. 1976. Organisation of the rosy locus in *Drosophila melanogaster*: Evidence for a control element adjacent to the xanthine dehydrogenase structural element. *Genetics*, *84*, 233-255.
- DAVIDSON, E. H., AND BRITTEN, R. J. 1973. Organisation, transcription and regulation in the animal genome. *Quart. Rev. Biol.*, *48*, 565-613.
- MATHER, K. 1953. Genetical control of stability in development. *Heredity*, *7*, 297-336.
- MATHER, K. 1975. Genotype  $\times$  environment interactions. II. Some genetical considerations. *Heredity*, *35*, 31-53.
- MATHER, K., AND HARRISON, B. J. 1949. The manifold effect of selection. *Heredity*, *3*, 1-52 and 131-162.
- SPICKETT, S. G. 1963. Genetic and developmental studies of a quantitative character. *Nature*, *199*, 870-873.
- THODAY, J. M. 1961. Location of polygenes. *Nature*, *191*, 368-370.
- WIGAN, L. G. 1949. The distribution of polygenic activity on the X chromosome of *Drosophila melanogaster*. *Heredity*, *3*, 53-66.