ESTIMATING THE COMPONENTS OF CONTINUOUS VARIATION

II. GENETICAL

PATRICIA COOKE and KENNETH MATHER
*Agricultural Research Council Unit of Biometrical Genetics,
Department of Genetics, University of Birmingham*

Received 1.x.61

1. INTRODUCTION

THE genetical components of continuous variation as estimated by biometrical analysis are compound: they properly include the effects of all the genes whose differences contribute to the variation of the character in question. Generally considerations of the components can be taken no further: the estimates must be accepted as such, subject to the standard errors that the nature of the experiment and the method of analysis make appropriate. In *Drosophila melanogaster*, however, the components of variation can also be approached from the opposite direction. The special stocks available make it possible to measure the effects of a given chromosome or part of a chromosome and to do so for virtually the whole of the genotype, so that values can be arrived at for the components of variation not only by biometrical analysis of the overall variation of the character but also by synthesis from the individual effects of such genetic units as it may be technically convenient to recognise.

An earlier paper (Cooke *et al.*, 1962) has reported the biometrical analysis of variation in the number of sternopleural chaetæ among the descendants of the cross between two inbred *Drosophila* lines, the Birmingham strain of Oregon (B) and Samarkand (S). Two methods of analysis were used and these results compared, viz. the simple unweighted analysis due to Mather (1949) and a more efficient, but more complex, weighted analysis described by Nelder (1960). The two methods yielded effectively the same estimates for the components of variation, though the weighted analysis showed certain advantages.

In the present paper, we are concerned with the opposite approach to the components of variation, by their synthesis from genetic differences measured individually, the three major chromosomes being taken as units for the purpose. Though obviously each chromosome is not a final unit in the determination of chaeta number, their use as such should give at least a first approximation to the synthetic values of the components.

2. MEASUREMENT OF THE INDIVIDUAL GENETIC EFFECTS

Since the three major chromosomes are being treated as units and the fourth chromosome neglected, eight stocks will suffice to provide all the possible true-breeding combinations of the genetic

elements to be followed in the experiment, and two of these eight will be the parent lines, B and S. Denoting chromosomes originating from the Oregon stock by B and those from the Samarkand stock by S, these stocks can be written as BBB and SSS respectively, the first letter indicating the X chromosome and the others the second and third chromosomes in turn. The six stocks required to complete the tally are thus BBS, BSB, BSS, SBB, SBS and SSB. These were made up by standard procedures from the parent stocks, using the *M5*, *Cy L4* and *Mé Sb* chromosomes as aids. Despite the inversions they contain, the marked chromosomes, especially *Mé Sb*, undergo some effective recombination with wild-type chromosomes, but the amount is sufficiently small for there to be little risk of the wild-type chromosomes as they appear in the final stocks, differing at all materially from those of the B and S lines with which the operation began.

A set of diallel crosses was then made up from the eight lines, males from each being crossed onto females from each to give 64 families in all. Of these 64, the 8 in which males and females were used from the same stock will of course produce the 8 true-breeding lines. In addition, however, all the possible combinations of heterozygotes for one or more chromosomes will be included among the remaining 56 families. The females include 27 genotypes in all and these occur among the 64 families in the classical tri-factorial F_2 proportions, *i.e.* the 8 homozygotes in 1 family each, the twelve single heterozygotes in 2 families each, the 6 double heterozygotes in 4 families each and the triple heterozygote in 8 families (see table 1). Since, however, one of the units is the X chromosome the males include only 18 genotypes falling into 2 groups of 9. One group carries the X from stock B and the other the X from stock S, the nine members of each group being the 9 possible combinations in respect of chromosomes II and III. The genotypes homozygous in chromosomes II and III as well as hemizygous in X are each represented by 2 families, the 8 heterozygous in respect of one autosome by 4 families each and the two doubly heterozygous genotypes by 8 families each (see table 1).

The sternopleural chaetæ were counted on 10 flies of each sex from each of the 64 families. All matings were between flies from true-breeding lines so that flies of like sex are of the same genotype within each family. Thus, apart from the sex difference, variation within families is non-genetic and estimation of the genetic parameters and their standard errors has been based therefore on the mean number of chaetæ of the 10 flies of each sex from each family. The mean numbers of chaetæ averaged over families of the same genotype are shown in table 1 for the 27 genotypes among females and the 18 genotypes among males. These overall genotype means provide the material for estimating the genetic parameters. The variation among the means of like genotype, not shown in the table, provide the estimates of error variation. The table also shows the number of

families which have contributed to each genotype mean. These are in fact double the numbers given above for the various genotypes because the experiment was duplicated. Since, however, no significant difference of any kind was found between the two blocks, their results have been pooled in the analysis as described here.

Taking females first, there are 128 family means distributed among 27 genotypes. The analysis will include therefore 127 degrees of freedom (d.f.) of which 26 are for differences between genotypes.

TABLE 1

Mean chaeta numbers and numbers of families of each genotype

II	BB			BS			SS		
III X	BB	BS	SS	BB	BS	SS	BB	BS	SS
Females									
BB	22.85 2	21.75 4	19.05 2	22.48 4	20.50 8	19.05 4	20.40 2	20.08 4	18.60 2
BS	22.05 4	21.19 8	20.28 4	21.84 8	20.47 16	19.36 8	21.35 4	19.69 8	18.60 4
SS	21.60 2	20.55 4	21.20 2	20.88 4	20.70 8	19.80 4	20.80 2	20.15 4	18.40 2
Males									
B	21.40 4	20.65 8	18.95 4	21.29 8	19.74 16	18.34 8	20.65 4	19.28 8	17.73 4
S	21.00 4	20.41 8	18.88 4	20.86 8	20.24 16	19.25 8	20.88 4	19.35 8	18.25 4

The remaining 101 d.f. consist of 26 for variation of the 26 genetic differences between blocks, 74 for differences among families within genotypes within blocks and 1 for the overall block difference. Since this last difference was found not to be significant it has been included with the 74. The sum of squares for this single d.f. is found in the obvious way as the square of the difference between the block totals (block total being the sum of the 64 family means included in that replicate of the experiment) divided by 128. The sum of squares for the 74 d.f. is found by the summation over all genotypes of the sums of squares of deviations of the mean of families of like genotype within the block, round their genotype mean within the block. Each block contributes in this way a sum of squares for 37 d.f. making 74 d.f. when the blocks are pooled.

The sum of squares for the 26 d.f. between genotypes is calculated from the 27 overall genotype means. Each mean is squared and multiplied by the number of families (varying from 2 to 16) which

has contributed to it. These quantities are then summed and the correction term for the whole experiment subtracted. The same operation can be carried out using the 27 genotype means for each block separately, the multipliers being reduced to half (*i.e.* varying from 1 to 8) and the correction term being that calculated from the block total. Two sums of squares are thus obtained from the two blocks, each for 26 d.f., which depend on the differences between genotypes within that block. Adding these gives a sum of squares for 52 d.f. which includes that for the 26 d.f. representing the differences between genotypes pooled over blocks and also a further item for 26 d.f. representing the variation of the genetic differences between

TABLE 2
Analysis of variance of differences among the mean chaeta numbers for females and males of the 128 families

Item	Females		Males	
	d.f.	M.S.	d.f.	M.S.
Differences among genotypes .	26	5.0002 *	17	6.6729 *
Heterogeneity of genotypes over blocks	26	0.4413	17	0.6200
Overall difference between blocks	1	0.1726	1	0.1512
Differences among families of like genotype within blocks	74	0.4173	92	0.4953
Pooled error variance	101	0.4211	110	0.5114

* = Very significant.

blocks. This latter item can thus be found as the difference between the S.S. for the 52 d.f. and that for the 26 d.f. representing overall genetic differences whose calculation was outlined above. The analysis of variance obtained in this way is set out in table 2.

The analysis of variance for the males is obtained in an exactly corresponding way, the difference being that with only 18 genotypes, there will be but 17 d.f. for genetic differences, 17 d.f. for variation of these differences between blocks, 92 ($= 46 \times 2$) for differences between families of like genotype within blocks and 1 for the overall block differences. The analysis is set out in table 2 alongside that for females.

The analyses of variance require little comment. In neither sex is the item for heterogeneity of genotypes (that is variation of genotypic differences over blocks) significant when tested against the mean square for differences among families of like genotype within blocks, and in both sexes the mean square for the overall difference between blocks is less than this same error variance. The three items were therefore pooled to give an error variance for $26+1+74 = 101$ d.f. in the case of females and for $17+1+92 = 110$ d.f. in the case of males. Whether tested against this pooled error variance or against

the mean square for differences within genotypes within blocks, the items for differences between genotypes is highly significant in both sexes as would indeed be expected.

The 26 d.f. for differences among the 27 female genotypes can be partitioned into 26 items each for 1 d.f. and each related to a specific genetic parameter. Using the notation of Hayman and Mather (1955), the 26 parameters comprise 3 d 's, depending on additive effects, one for each chromosome; 3 h 's, depending on dominance, one for each chromosome; 3 i 's, depending on first order interactions of the additive effects (*i.e.* $d \times d$ interactions); 6 j 's depending on first order interactions of additive and dominance effects (*i.e.* $d \times h$ interactions); 3 l 's, depending on first order interactions of dominance effects (*i.e.* $h \times h$ interactions); 1 $d \times d \times d$ second order interaction of additive effects; 3 $d \times d \times h$ second order mixed interactions; 3 $d \times h \times h$ second order mixed interactions; and 1 $h \times h \times h$ second order interaction of dominance effects. Following Hayman and Mather's definition of the first order interactions and extending them to those of the second order, these 26 parameters may be estimated from orthogonal comparisons among the 27 genotypes. When so estimated they will provide a complete description of the genetic differences and will also contribute, independently of one another, to the variation in chaetæ number in the F_2 between the two original inbred lines.

The 17 genetic comparisons among the 18 male genotypes can similarly be used to estimate the 17 genetic parameters analogous to the 26 of the females, but omitting h for the X chromosome and also all the interactions which include this h , since in males the X is always hemizygous so that no properties of dominance can be ascribed to it. The genetic parameters of both sexes are listed in table 3.

Estimates of the parameters can be obtained from the data of table 1 though it should be noted that they are derived not directly from the genotype means as shown in that table but from the genotypic totals, *i.e.* from the products of the mean and the number of families from which the mean is derived. Thus the BB BB BB female genotype has a mean of 22.85 based on 2 families so that its total is 45.70, the triple heterozygote has a mean of 20.46875 based on 16 families so that its total is 327.50. The d_1 and h_1 for the X chromosomes are then found from the totals obtained by separately summing BB, BS and SS for the X over all genotypes in respect of the two autosomes. These three values are obtained by summing along the three rows of table 1 [*e.g.* that for BB being $(22.85 \times 2) + (21.75 \times 4) + \dots + (18.60 \times 2)$]. Then d_1 is given as the difference between the two homozygote sums, divided by the total number of families involved *i.e.* 32 for each homozygote making 64 in all. We thus find d_1 as

$$\frac{1}{64}[S(\text{BB}, \bullet, \bullet) - S(\text{SS}, \bullet, \bullet)] = \frac{1}{64}[659.20 - 655.10] = 0.0641$$

where $S(\text{BB}, \bullet, \bullet)$ stands for the sum over all families homozygous

for the X chromosome from the B stock. In all calculations the S values are always subtracted from the B values so that a positive value for d means that the B chromosomes are making the greater contribution to the chaeta number. The sign of h is dependent on the

TABLE 3
The values of the 26 genetic parameters in females and the 17 in males

	Females		Males	
	Parameter	SS	Parameter	SS
d_1	0.0641	0.2627	-0.1234	1.9503
d_2	0.6687	28.6225**	0.5375	18.4900**
d_3	1.1359	82.5827**	1.1407	83.2656**
$\frac{1}{2}h_1$	-0.0086	0.0095	...	...
$\frac{1}{2}h_2$	0.0273	0.0957	0.0406	0.2113
$\frac{1}{2}h_3$	0.0070	0.0063	0.0344	0.1513
i_{12}	0.1156	0.4278	0.0531	0.1806
i_{13}	0.4687	7.0313**	0.2688	4.6225**
i_{23}	-0.1218	0.4753	-0.2469	1.9503
$\frac{1}{2}j_{12}$	-0.0078	0.0039	-0.0641	0.5253
$\frac{1}{2}j_{21}$	0.0031	0.0006	...	...
$\frac{1}{2}j_{13}$	0.0265	0.0452	0.0172	0.0378
$\frac{1}{2}j_{31}$	0.0485	0.1502	...	...
$\frac{1}{2}j_{23}$	-0.0343	0.0756	0.0719	0.3306
$\frac{1}{2}j_{32}$	0.0453	0.1314	0.0000	0.0000
$\frac{1}{4}l_{12}$	-0.0117	0.0176	...	...
$\frac{1}{4}l_{13}$	-0.0726	0.6757	...	...
$\frac{1}{4}l_{23}$	-0.0273	0.0957	-0.0063	0.0050
$d_1d_2d_3$	0.5000	4.0000**	0.1281	0.5253
$\frac{1}{2}d_1d_2h_3$	0.2031	1.3203	0.0250	0.0400
$\frac{1}{2}d_1h_2d_3$	0.1187	0.4513	0.0656	0.2756
$\frac{1}{2}h_1d_2d_3$	-0.1218	0.4753	...	...
$\frac{1}{4}d_1h_2h_3$	-0.1828	2.1389*	-0.0828	0.8778
$\frac{1}{4}h_1d_2h_3$	0.1125	0.8100	...	...
$\frac{1}{4}h_1h_2d_3$	0.0078	0.0039	...	...
$\frac{1}{8}h_1h_2h_3$	0.0273	0.0957	...	...

* Significant at the 5 per cent. level.

** Significant at the 1 per cent. level.

direction of dominance, *i.e.* a negative h indicates that dominance is in the direction of the lower chaeta number irrespective of which of the two parents makes the higher contribution.

Similarly since there are 64 BS families as compared with 32 of each homozygote

$$\begin{aligned}\frac{1}{2}h_1 &= \frac{1}{128} [S(\text{BS}, \bullet, \bullet) - S(\text{BB}, \bullet, \bullet) - S(\text{SS}, \bullet, \bullet)] \\ &= \frac{1}{128} [1313.20 - 659.20 - 655.10] = -0.0086\end{aligned}$$

or $h_1 = -0.0172$.

The primary estimate is of $\frac{1}{2}h_1$, not of h_1 itself, since the orthogonal comparisons being used provide estimates of the parameters as they contribute to deviations from the mean. Half the observations on the females in the experiment are of flies heterozygous (BS) for the X chromosome so that the overall mean is $\frac{1}{2}h_1$ in respect of the effect of the sex chromosome. Thus the deviation of these BS flies from the mean is $h_1 - \frac{1}{2}h_1 = \frac{1}{2}h_1$ which is the quantity estimated. The other d 's and h 's are found similarly.

The first order interaction parameters between, for example, chromosomes X and II are found from the genotype totals for the nine genotypes in respect of chromosomes X and II, the summation being carried out for each of the nine over the three corresponding genotypes in respect of chromosome III. The totals thus obtained for the 9 genotypes in respect of chromosomes X and II are given in table 4. The marginal sums of this table provide, of course, the data for evaluating d_1 , h_1 , d_2 and h_2 by the procedure already elaborated.

Referring to Hayman and Mather's table 1 it is clear that i_{12} , the interaction of d_1 and d_2 , will be found as

$$i_{12} = \frac{1}{32}[S(\text{BB}, \text{BB}, \bullet) + S(\text{SS}, \text{SS}, \bullet) - S(\text{BB}, \text{SS}, \bullet) - S(\text{SS}, \text{BB}, \bullet)] \\ = \frac{1}{32}[170.80 + 159.00 - 158.30 - 167.80] = 0.1156$$

the divisor being 32 because only those families homozygous in respect of both chromosomes enter into the calculation.

Similarly with the $d \times h$ interactions,

$$\frac{1}{2}j_{12} = \frac{1}{64}[S(\text{SS}, \text{BB}, \bullet) + S(\text{SS}, \text{SS}, \bullet) - S(\text{SS}, \text{BS}, \bullet) - S(\text{BB}, \text{BB}, \bullet) \\ - S(\text{BB}, \text{SS}, \bullet) + S(\text{BB}, \text{BS}, \bullet)] \\ = \frac{1}{64}[167.80 + 159.00 - 328.30 - 170.80 - 158.30 + 330.10] \\ = -0.0078,$$

and

$$\frac{1}{2}j_{21} = \frac{1}{64}[S(\text{BB}, \text{SS}, \bullet) + S(\text{SS}, \text{SS}, \bullet) - S(\text{BS}, \text{SS}, \bullet) \\ - S(\text{BB}, \text{BB}, \bullet) - S(\text{SS}, \text{BB}, \bullet) + S(\text{BS}, \text{BB}, \bullet)] \\ = \frac{1}{64}[158.30 + 159.00 - 317.30 - 170.80 - 167.80 + 338.80] \\ = 0.0031$$

the quantities estimated being $\frac{1}{2}j_{12}$ and $\frac{1}{2}j_{21}$ as reference to Hayman and Mather's table makes clear.

Finally, taking the $h \times h$ interaction,

$$\frac{1}{4}l_{12} = \frac{1}{128}[S(\text{BB}, \text{BB}, \bullet) + S(\text{BB}, \text{SS}, \bullet) - S(\text{BB}, \text{BS}, \bullet) \\ + S(\text{SS}, \text{BB}, \bullet) + S(\text{SS}, \text{SS}, \bullet) - S(\text{SS}, \text{BS}, \bullet) \\ - S(\text{BS}, \text{BB}, \bullet) - S(\text{BS}, \text{SS}, \bullet) + S(\text{BS}, \text{BS}, \bullet)] \\ = \frac{1}{128}[170.80 + 158.30 - 330.10 + 167.80 + 159.00 - 328.30 \\ - 338.80 - 317.30 + 657.10] = -0.0117.$$

The first order interactions of the effects of chromosomes X and III and II and III are found similarly.

The calculation of the 8 second order, or three chromosome, interactions involves the use of the 27 genotype sums individually. It

may be illustrated by reference to the calculation of the simplest of the 8 parameters, that measuring the $d \times d \times d$ interaction. This involves only the sums of the 8 fully homozygous genotypes as

$$\begin{aligned} & \frac{1}{16}[S(\text{BB, BB, BB}) + S(\text{SS, SS, BB}) + S(\text{SS, BB, SS}) \\ & \quad + S(\text{BB, SS, SS}) - S(\text{BB, BB, SS}) - S(\text{BB, SS, BB}) \\ & \quad - S(\text{SS, BB, BB}) - S(\text{SS, SS, SS})] \\ &= \frac{1}{16}[45 \cdot 70 + 41 \cdot 60 + 42 \cdot 40 + 37 \cdot 20 - 38 \cdot 10 - 40 \cdot 80 - 43 \cdot 20 - 36 \cdot 80] \\ &= 0 \cdot 5000, \end{aligned}$$

TABLE 4

The data of table 1 rearranged for the calculation of interactions between chromosomes X and III. The entries are chaeta totals, not means as in table 1. The numbers of families over which summation has proceeded are shown in brackets

<div>II X</div>	BB	BS	SS	Total
Females				
BB	170.8 (8)	330.1 (16)	158.3 (8)	659.2 (32)
BS	338.8 (16)	657.1 (32)	317.3 (16)	1313.2 (64)
SS	167.8 (8)	328.3 (16)	159.0 (8)	655.1 (32)
Total	677.4 (32)	1315.5 (64)	634.6 (32)	2627.5 (128)
Males				
B	326.6 (16)	632.8 (32)	307.7 (16)	1267.1 (64)
S	326.8 (16)	644.8 (32)	311.3 (16)	1282.9 (64)
Total	653.4 (32)	1277.6 (64)	619.0 (32)	2550.0 (128)

the divisor being 16 as only 16 of the 128 families contribute to the estimate. The functions for estimating the remaining seven parameters are easily found by applying the principles of orthogonal functions as described by, for example, Fisher (1953) or Mather (1943).

The estimates thus obtained for the 26 genetical parameters of the females and similarly for the 17 parameters of the males are set out in table 3. The significance of their contributions to the differences among the 27 female and 18 male genotypes is readily tested by calculating the contribution each parameter makes to the sum of squares for differences among the genotypes. To take d_1 in females as an example, we note that it was estimated as

$$\frac{1}{64}[S(\text{BB, } \bullet, \bullet) - S(\text{SS, } \bullet, \bullet)],$$

so that its contribution to the sum of squares is

$$\begin{aligned} SS(d_1) &= \frac{1}{64} [S(BB, \bullet, \bullet) - S(SS, \bullet, \bullet)]^2 \\ &= \frac{1}{64} [659 \cdot 20 - 655 \cdot 10]^2 = 0 \cdot 2627, \end{aligned}$$

as shown in the second column of table 3.

Similarly we found

$$\frac{1}{2}h_1 = \frac{1}{128} [S(BS, \bullet, \bullet) - S(BB, \bullet, \bullet) - S(SS, \bullet, \bullet)]$$

so that the contribution to the sum of squares is

$$\begin{aligned} SS(\frac{1}{2}h_1) &= \frac{1}{128} [S(BS, \bullet, \bullet) - S(BB, \bullet, \bullet) - S(SS, \bullet, \bullet)]^2 \\ &= \frac{1}{128} [1313 \cdot 20 - 659 \cdot 20 - 655 \cdot 10]^2 = 0 \cdot 0095. \end{aligned}$$

The remaining sums of squares given in table 3 are found in the same way. Each of the sums of squares corresponds to 1 d.f. and it will be observed that the 26 from females add up to give the sum of squares already found by a different method for the 26 degrees of freedom between genotypes in table 2. The total of the 17 items for the males agrees similarly with the corresponding total sum of squares for 17 d.f. in the analysis of variance for males in table 2.

Each of the contributions made by the individual parameters to the sum of squares can thus be tested by comparison with the error variance, for females or males as the case may be, already found in table 2. The error variances are based on large numbers of degrees of freedom. If the sums of squares in columns 2 and 4 of table 3 are divided by the appropriate error variance, the ratio may be regarded as a χ^2 for 1 d.f. and the corresponding probability obtained from the table of χ^2 , so providing the test of significance for the contribution of that parameter to the differences among genotypes. The assessments of significance made in this way are shown in the table.

Before proceeding to an examination of this table it should also be observed that, even apart from any general effect of the sex difference, the parameters as estimated from the two sexes are not necessarily alike. Thus, if we refer to Hayman and Mather's table describing the relations of the nine genotypes produced by two segregating genes in terms of the eight genetical parameters, j_{ba} for example can be isolated and estimated independently of the other seven and in particular independently of d_b . Since our female data correspond to an F_2 , we can and do estimate j_{21} and j_{31} which correspond to j_{ba} . But if $A-a$ represents a sex-linked gene, Aa genotypes are impossible in males. It will thus be seen from the AA and aa column of Hayman and Mather's table that d_b and $\frac{1}{2}j_{ba}$ become inseparable and that the comparison of BB with bb homozygotes, summed over the two classes in respect of $A-a$, must yield an estimate of $(d_b - \frac{1}{2}j_{ba})$. This in fact is the estimate appearing as d_2 in males, when expressed in terms of the female parameters. Similarly the d_3 of males is $(d_3 - \frac{1}{2}j_{31})$ of females and h_2 and h_3 of males are $(h - \frac{1}{2}l_{12})$ and $(h_3 - \frac{1}{2}l_{13})$ of females. Higher order interactions will also be involved in a similar way. Since,

however, the j , l and corresponding second order interactions are in general small and insignificant, this theoretical complication need not worry us in the practical comparisons of the main effects as we have measured them from the different sexes.

The results of the analysis are clear. The d 's are highly significant for both autosomes in both sexes and the effect of chromosome III is consistently about twice that of chromosome II. There is no clear indication of a similar effect of chromosome X, but this is a point we will take up again later. No h is significant in either sex, so that there is no reliable evidence of dominance from this analysis.

The correspondence of i_{13} from both sexes provides clear evidence of first order interaction between the d effects of chromosomes X and III and there is a further indication from the females of a second order interaction between the d effects of all three chromosomes, though this lacks support from the males. The apparent second-order interaction of $h_1 d_2 h_3$ in females should be treated with reserve: it is significant only at the 5 per cent. level, it is not supported by the males; and out of so many parameters one must expect an occasional case of apparent significance at this level arising solely by sampling variation.

The first order interaction, i_{13} , between the d effects of chromosomes X and III, is perhaps a little suspicious at first sight since d_1 is not significant in either sex. There is, however, another way of assessing the value of d_1 , by comparison of the males in the F_1 's and the females of the F_2 's of the reciprocal crosses between the original stocks B and S. The mean chaeta numbers are given for both sexes in table 5, where expectations in terms of d_1 , d_2 , d_3 , h_1 , h_2 and h_3 are also shown. The F_1 females have the same expectations in the two reciprocal crosses and indeed their average chaeta number differs only by $20.4586 - 20.3138 = 0.1448$. Equally the F_2 males have the same expectation in the two reciprocals and they differ only by $19.7091 - 19.8228 = -0.1133$, the sign indicating that those for the $S \times B$ cross have the larger value. But the males of F_1 do not have the same expectation in the reciprocals and if we subtract the value for the $S \times B$ males from that for $B \times S$ males, we find $2d_1 = -0.5344$ or $d_1 = -0.2672$, the sign indicating that the X chromosome for S contributes a higher chaeta number than that of B. This estimate of d_1 is biased by j_{12} which is confounded with it; but the value of j_{12} is so small (table 3), that the bias must be negligible. In the same way the difference between the females in the F_2 's of the $B \times S$ and $S \times B$ crosses yield $d_1 = -0.3387$, which is unbiased by confounding with any interaction.

Both sexes of the BIP generations yield unbiased estimates, that from males being $d_1 = -0.3557$ and that from females $d_1 = -0.5956$. All four estimates from the F_1 's, F_2 's and BIP's are consistent in sign and reasonably consistent in magnitude: they all materially exceed the two values found for differences which must be attributed to error

variation. Furthermore the d_1 from males in table 3 is also negative and only just falls short of significance at the 5 per cent. level. The female d_1 in that table is positive but nowhere near approaches significance. If we take all six estimates of d_1 which are now available, their mean is -0.2694 and the standard error of this mean calculated from the variation among the six estimates themselves is 0.0915 , giving $t_{[5]} = 2.94$ with a probability between 0.05 and 0.02 . There can thus be little doubt that chromosome X is exerting a d effect so

TABLE 5

Mean number of chaeta and expected deviation from the mid-parent in terms of d and h of the parents, F_1 's, F_2 's and BIP's of the reciprocal crosses. The d 's are all arbitrarily taken as positive in the B parent

Generation	Males	Females
Parents		
B . .	22.4724 $d_1 + d_2 + d_3$	22.9310 $d_1 + d_2 + d_3$
S . .	18.2034 $-d_1 - d_2 - d_3$	18.5207 $-d_1 - d_2 - d_3$
F_1		
B $\times$ S .	19.5828 $d_1 + h_2 + h_3$	20.4586 $h_1 + h_2 + h_3$
S $\times$ B .	20.1172 $-d_1 + h_2 + h_3$	20.3138 $h_1 + h_2 + h_3$
F_2		
B $\times$ S .	19.7091 $\frac{1}{2}(h_2 + h_3)$	20.1205 $\frac{1}{2}d_1 + \frac{1}{2}(h_1 + h_2 + h_3)$
S $\times$ B .	19.8224 $\frac{1}{2}(h_2 + h_3)$	20.4592 $-\frac{1}{2}d_1 + \frac{1}{2}(h_1 + h_2 + h_3)$
BIP		
B $\times$ S .	19.4841 $\frac{1}{2}d_1 + \frac{1}{2}(h_2 + h_3)$	20.0607 $\frac{1}{4}d_1 + \frac{1}{2}(h_1 + h_2 + h_3)$
S $\times$ B .	19.8398 $-\frac{1}{2}d_1 + \frac{1}{2}(h_2 + h_3)$	20.3585 $-\frac{1}{4}d_1 + \frac{1}{2}(h_1 + h_2 + h_3)$

that the significant i_{13} is not in fact surprising. We may note that this d effect of chromosome X is only about half that of chromosome II and one-quarter of that of chromosome III. It is of opposite sign to d_2 and d_3 showing that the X chromosome from stock S has a greater chaeta producing ability than that of the X chromosome from B, whereas with both autosomes B carries the chromosomes of higher effect.

3. THE SYNTHETIC COMPONENTS OF VARIATION

The values expected for $D = S(d^2)$ and $H = S(h^2)$ can be found from the estimates of the parameters in table 3. Thus in females $D = 0.0641^2 + 0.6687^2 + 1.1359^2 = 1.6053$. These synthetic values are

set out in table 6 where they are compared with the estimates obtained by Cooke *et al.* by biometrical analysis using both the simple unweighted and the iterative weighted procedures. Both the synthetic and the analytical values of D and H are given for the sexes separately. Only a single analytical value is given for males since the results from the reciprocal crosses $B \times S$ and $S \times B$ were fully consistent in this sex. The results for females were not, however, consistent over reciprocals and indeed those for the $B \times S$ females showed internal evidence of aberrancy. In addition, therefore, to the estimates from the females pooled over reciprocals, the table also gives those for the $S \times B$ reciprocal

TABLE 6

Comparison of the analytic and synthetic estimates of D and H

Component		Analysis		Synthesis (data from table 3)		
		Unweighted	Weighted			
D	Males	2.69±0.98	2.39±0.31	}	1.61	
	Females { All . . . { SB . . .	3.41±0.25	3.38±0.33		}	1.74
		2.96±0.24	2.75±0.46			
H	Males	2.59±2.41	1.49±0.77	}	0.01	
	Females { All . . . { SB . . .	-0.68±0.63	-0.69±0.73		}	0.00
		3.01±0.59	2.55±1.00			

alone, the results from these females being not only internally consistent, but also consistent with both sets of males. The standard errors attached to the unweighted estimates should be treated with caution since they are based on only one degree of freedom (see Cooke *et al.*).

Two features of this table are striking. In the first place, the synthetic values of D are only about 60 per cent. of the analytical values. Secondly, while the synthetic values for H indicate that dominance is very low, if it is present at all, the analytical values suggest the contrary. The evidence of the analytical values should not be given too much weight in respect of H: even if we ignore the insignificant negative values for H obtained from the combined females as due to the aberrancy of the $B \times S$ females, the standard errors for the positive values remaining from the $S \times B$ females and the pooled males are high and only one entry, from the weighted analysis of $S \times B$ females, could be regarded as significant. Nevertheless, the synthetic H values do appear to fall short of the analytical, just as with D and possibly to an even greater extent. Several factors might contribute to the creation of these differences. The fourth chromosome has been neglected: the analytical values were obtained

assuming autosomal inheritance of all the genes, an assumption which is obviously wrong in respect of any genes on the X chromosome: the synthetic values explicitly exclude interactions, among which i_{13} is notably significant, whereas the analytical values can hardly be uninfluenced by such an interaction: the values in table 3 have been used for d_1 , whereas we have seen above that there is reason to believe d_1 to be greater than this table would suggest: some difference between the B and S chromosomes may have been lost during the process of building up the eight lines from which the estimates used in the synthetic D were obtained: and finally the chromosomes were treated as units in obtaining the estimates of the genetic parameters used in the synthesis, so that these will fail to reflect any consequence of recombination in releasing variation in the segregating generations whose results provided the material for the biometrical analysis.

Several of these factors can probably be ruled out as making but minor contributions at most to the differences between the analytical and synthetic values of D. In the first place, neglect of the fourth chromosome could hardly be regarded as responsible for a 40 per cent. reduction in the expected variation. Next, despite its significance, i_{13} would add only 0.21 to the female D and less than 0.1 to the male D if its square were included in its entirety as a separate item, and if all three i 's were so included they would jointly add but 0.25 to the female D and 0.14 to the male D. Furthermore, the inclusion of interactions as they stand in table 3 would hardly affect the value of H, since the j and l interactions are small. The underestimation of d_1 would also have but a minor effect: using the average value of 0.29 obtained from the F_1 , F_2 and BIP means in place of the figure from table 3, would raise the female D to only 1.81 and the male D to only 1.66, and should there be any similar underestimation of h there is no reason to believe that it would account for any greater fraction of the difference in H. Also, in the case of the inversions used, any loss of difference between the B and S chromosomes in the stock building programme is hardly likely to be so great as to account for the differences observed in either D or H. The two factors to consider further, therefore, are the assumptions of autosomal inheritance and the possible effects of recombination.

While it is difficult to assess the effects of sex-linkage on the values obtained for D and H yielded by an analysis which assumes all inheritance to be autosomal, the matter can be approached in a different way by synthesising from the estimates of the genetic parameters in table 3, not D and H but the values expected for the statistics V_{1F2} , V_{1S2} , V_{2S3} and W_{1S23} , from whose interrelations D and H are derived biometrically. Allowance can thus be made for sex-linkage in deriving the expectations for these statistics in terms of the genetic parameters. The formulæ for V_{1F2} in males and females are given at the top of table 7. All the first order interactions are included but the second order interactions have been omitted for simplicity.

The males have the same V_{1F2} expectation no matter from which reciprocal cross they come, but the female formulæ differ in some terms according to the direction of the cross. This is the reason for the $\pm$ and $\mp$ signs in certain of the terms, the upper of the two gives the sign appropriate to the $S \times B$ cross and the lower to $B \times S$, since the X chromosome for S has the higher chaeta producing activity. It should be noted that, as observed in the previous section, the d_2 and d_3 of males correspond to $(d_2 - \frac{1}{2}j_{21})$ and $(d_3 - \frac{1}{2}j_{31})$ of females and the h_2 and h_3 of males to $(h_2 - \frac{1}{2}l_{12})$ and $(h_3 - \frac{1}{2}l_{13})$ of females.

TABLE 7

Formulation of the F_2 and BIP statistics in terms of the effects of the X, II and III chromosomes. Only second order interactions are excluded from the formulation of V_{1F2} , but only the d and i parameters are included in the formulation of the other statistics

$$\begin{aligned}
 &V_{1F2} \\
 &\text{Females} = \frac{1}{4}(d_1 \mp h_1)^2 + \frac{1}{4}d_2^2 + \frac{1}{4}(d_2 \pm i_{12})^2 + \frac{1}{4}d_3^2 + \frac{1}{4}(d_3 \pm i_{13})^2 + \frac{1}{8}h_2^2 + \frac{1}{8}(h_2 \pm j_{12})^2 \\
 &\quad + \frac{1}{8}h_3^2 + \frac{1}{8}(h_3 \pm j_{13})^2 + \frac{1}{8}i_{23}^2 + \frac{1}{8}j_{21}(j_{21} \mp 2i_{12}) - \frac{1}{8}j_{31}(j_{31} \mp 2i_{13}) \\
 &\quad + \frac{1}{16}l_{12}(l_{12} \mp 2j_{12}) + \frac{1}{16}l_{13}(l_{13} \mp 2j_{13}) + \frac{1}{8}j_{23}^2 + \frac{1}{8}j_{32}^2 + \frac{1}{16}l_{23}^2 + E_1 \\
 &\text{Males} = d_1^2 + \frac{1}{2}(d_2 - \frac{1}{2}j_{21})^2 + \frac{1}{2}(d_3 - \frac{1}{2}j_{31})^2 + \frac{1}{4}(h_2 - \frac{1}{2}l_{12})^2 + \frac{1}{4}(h_3 - \frac{1}{2}l_{13})^2 \\
 &\quad + \frac{1}{2}i_{12}^2 + \frac{1}{2}i_{13}^2 + \frac{1}{4}i_{23}^2 + \frac{1}{4}j_{12}^2 + \frac{1}{4}j_{13}^2 + \frac{1}{8}j_{23}^2 + \frac{1}{8}j_{32}^2 + \frac{1}{16}l_{23}^2 + E_1 \\
 &V_{1S3} \\
 &\text{Females} = \frac{5}{16}d_1^2 + \frac{3}{16}d_2^2 + \frac{3}{16}d_3^2 + \frac{3}{32}i_{12}^2 + \frac{3}{32}i_{13}^2 + \frac{1}{16}i_{23}^2 + \frac{1}{16}(d_2 \pm i_{12})^2 + \frac{1}{16}(d_3 \pm i_{13})^2 + E_2 \\
 &\text{Males} = \frac{1}{2}d_1^2 + \frac{1}{8}d_2^2 + \frac{1}{8}d_3^2 + \frac{1}{16}i_{23}^2 + \frac{1}{8}(d_2 \pm i_{12})^2 + \frac{1}{8}(d_3 \pm i_{13})^2 + E_2 \\
 &V_{2S3} \\
 &\text{Females} = \frac{1}{8}d_1^2 + \frac{3}{16}d_2^2 + \frac{3}{16}d_3^2 + \frac{3}{32}i_{12}^2 + \frac{3}{32}i_{13}^2 + \frac{1}{16}i_{23}^2 + \frac{1}{16}(d_2 \pm i_{12})^2 + \frac{1}{16}(d_3 \pm i_{13})^2 + E_1 \\
 &\text{Males} = \frac{1}{2}d_1^2 + \frac{1}{8}d_2^2 + \frac{1}{8}d_3^2 + \frac{1}{4}i_{12}^2 + \frac{1}{4}i_{13}^2 + \frac{1}{16}i_{23}^2 + \frac{1}{8}(d_2 \pm i_{12})^2 + \frac{1}{8}(d_3 \pm i_{13})^2 + E_1 \\
 &W_{1S23} \\
 &\text{Females} = \frac{5}{16}d_1^2 + \frac{3}{16}d_2^2 + \frac{3}{16}d_3^2 + \frac{3}{32}i_{12}^2 + \frac{3}{32}i_{13}^2 + \frac{1}{16}i_{23}^2 + \frac{1}{16}(d_2 \pm i_{12})^2 + \frac{1}{16}(d_3 \pm i_{13})^2 \\
 &\text{Males} = \frac{1}{2}d_1^2 + \frac{1}{4}d_2^2 + \frac{1}{4}d_3^2 + \frac{1}{16}i_{12}^2 + \frac{1}{16}i_{13}^2 + \frac{1}{16}i_{23}^2 \pm \frac{1}{16}d_2i_{12} \pm \frac{1}{16}d_3i_{13}
 \end{aligned}$$

The remaining formulæ shown in table 7 neglect not only the second order interactions but also the h , j and l parameters. These were all found to have only small values, so that their omission should have but little effect on the arithmetic comparison, while their inclusion in the formulæ would have added formidably to the algebra. In all cases there is a difference according to the direction of the initial cross as indicated by the inclusion of $\pm$ in appropriate terms just as with the V_{1F2} of females. The formulations are completed by the addition of the appropriate component of environmental variation, E_1 in the cases of V_{1F2} and V_{2S3} , E_2 in V_{1S3} , and no addition at all in W_{1S23} .

The expected or synthetic values have been obtained by substituting in these formulæ the estimates of the genetical parameters for the appropriate sex from table 3. The E_1 and E_2 values were those used by Cooke *et al.* They were obtained from the parental and F_1 generations, the sexes being taken separately since males are rather more variable than females. The synthetic values so obtained for the statistics are compared with those observed by Cooke *et al.*

in table 8. As we have had reason to observe earlier, the statistics for the $B \times S$ females are somewhat aberrant, the values of V_{1F2} and V_{2S3} being very low and that for W_{1S23} very high as compared both with one another and with the corresponding statistics from the other parts of the experiment. This appears in table 8 as a great excess of the value observed over that expected for W_{1S23} in the females and a much smaller excess (though still an excess) of the observed V_{1F2} and V_{2S3} over expectation as compared with the $S \times B$ females and both sets of males. Leaving aside the $B \times S$ females the remaining results are striking in the story they tell. The values expected always fall short of those observed and they do so by a proportional amount

TABLE 8

Values observed and expected for the F_2 and BIP statistics. The values taken for the environmental components are shown at the foot of the table

Statistic	Females				Males			
	$B \times S$		$S \times B$		$B \times S$		$S \times B$	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
$V_{1F2} (E_1)$	3.15	2.84	4.17	3.45	4.86	3.55	4.43	3.35
$V_{1S3} (E_2)$	1.19	0.77	1.20	0.92	1.07	0.74	1.20	0.91
W_{1S23}	1.05	0.38	0.74	0.53	0.58	0.35	0.54	0.47
$V_{2S3} (E_1)$	2.87	2.59	3.13	2.74	3.28	2.84	3.31	3.00

$$E_1 = 2.19$$

$$E_2 = 0.39$$

$$E_1 = 2.47$$

$$E_2 = 0.41$$

which is as consistent as could be expected for each statistic over the three sets of data, but which varies in an informative way between the statistics. The consistency for each statistic over the sets of data emphasises a point which is indeed apparent from inspection of the table, that the differences sex linkage leads us to expect between sexes and perhaps more particularly between reciprocals are in fact manifested in general by the observed values. The differences between statistics bring out that it must be the genetical component by which expectation chiefly falls short of observation. Averaging over the three sets of data, expectation is 73 per cent. of the value observed for W_{1S23} , 74 per cent. for V_{1S3} , 77 per cent. for V_{1F2} and 88 per cent. for V_{2S3} . Thus W_{1S23} , the statistic with the lowest non-heritable component (in fact no such component) shows the greatest excess of observation over expectation and V_{2S3} , which has proportionally the greatest non-heritable component, shows the lowest excess over expectation. The excess of observed over expected variation is thus clearly related to the genetic components and not to the non-heritable.

The genetical components can be found for both expected and observed values of each statistic by subtracting the appropriate E from the figures given in table 8. When this is done the genetical values expected can be expressed as percentages of the observed. These are not completely consistent over either statistics or sets of data: the proportion is lowest in V_{1F2} and highest in W_{1S23} , and it is lower from $B \times S$ males than either sex of $S \times B$. If, however, we average the proportions over all statistics for the three sets of data we obtain a proportion of 0.59 for the expected of the observed and this compares extremely closely with the average ratio the expected synthetic D 's bear to the D 's obtained by biometrical analysis, viz. 0.67 if we use the estimate for the weighted analysis and 0.60 if we use the unweighted figures (table 6). It is clear therefore that the shortfall of the synthetic components relative to those obtained by biometrical analysis is not due to the failure of the analysis to take sex-linkage into account. This shortfall is as clear and as large in the statistics as in the analytical D : its cause must be sought elsewhere and we will now examine the remaining possibility, that it springs from the effects of recombination releasing variability hidden in balanced combinations within the chromosomes of the original stocks B and S . Before we do so, however, it is worth noting that the consistency of the results from the statistics and the components encourages the belief that the failure to take sex-linkage into account has not jeopardised the results of a biometrical analysis, at least when the sex-linked elements are not of large effect.

4. THE EFFECTS OF RECOMBINATION

Tests of the consequences for the value of D of recombination in releasing variability have been carried out for chromosomes X and III . The method used was similar for the two chromosomes and will be illustrated by reference to the X chromosome. A cross was made between the lines BBB and SBB , and the M_5 chromosome was used to extract 20 X chromosomes from the progeny of the heterozygous females. Each of these chromosomes was then incorporated into a true breeding line whose constitution may be designated as RBB , where R stands for a chromosome extracted from a heterozygous female and hence prospectively recombinant. Similarly 20 chromosomes were extracted from the cross BSS and SSS , to give 20 lines of type RSS . This group differs essentially from the other in that it has a background of S as opposed to B chromosomes. Cultures were raised of each of these lines and the mean chaeta numbers determined from counts of 10 males and 10 females. The variances of these means among the 20 RBB lines and again among the 20 RSS lines give measures of the genetical variation among a sample of X chromosomes displaying the effects of recombination. They can be compared with the variances of means within groups consisting of BBB and SBB lines and BSS and SSS lines respectively. It was intended to raise five

cultures of each of these four lines, but while five were successful in BBB and BSS, only four each were obtained in SBB and SSS.

The results of these tests are set out in table 9a and b. It will be seen from the former table that the overall mean chaeta number of the recombinant group closely matches the average of the B and S

TABLE 9

Results of X chromosome recombination

(a) Mean chaeta number with parental and recombinant X chromosomes

Background	Sex	Recombinant	Parents		Mean parent
			B	S	
B	Males . .	21.51	21.64	21.80	21.72
	Females . .	22.21	23.02	22.22	22.62
S	Males . .	18.16	17.62	17.65	17.64
	Females . .	18.63	17.86	18.18	18.02

(b) Variation among parental and recombinant cultures

Background	Sex	Mean squares		Significance		
		Recombinants 19 d.f.	Parental 8 d.f.	Error		
				Own V.R.	Pooled V.R.	Previous χ^2 [19]
B	Males . .	1.618	0.031	52.04***	9.37***	62.08***
	Females . .	1.208	0.410	2.95	6.99***	55.02***
	Joint . .	1.237	0.081	15.26***	18.87***	
S	Males . .	0.827	0.128	6.48**	4.78***	31.72*
	Females . .	0.524	0.123	4.28*	3.03**	23.87
	Joint . .	0.593	0.050	11.85***	9.04***	

* Significant at the 5 per cent. level.

** Significant at the 1 per cent. level.

*** Significant at the 0.1 per cent. level.

parents on the B background but is slightly higher than the parental average on the S background. The parental difference is of course small in both cases. Turning to table 9b, it will be seen that on both backgrounds both sexes show a higher variance between means of the 20 lines in the recombination group than between the means of the 9 cultures in the parental group, the variance being greater in the recombinant group on the B background than on the S but barely significantly so. In three of the four cases distinguished by sex and

background, the excess of the recombinant variance is significant when tested against its own corresponding parental variance. The significance of the excess is even more striking if we pool the 10 males and 10 females to find a joint mean for each culture, recombinant and parental. The parental variances for the sex means themselves vary somewhat but they are not significantly heterogeneous. The four have therefore been pooled to give a pooled parental variance of 0.1728 for 36 d.f., of which 9 are contributed by each sex on each background. Equally the parental variances for the joint means can be pooled over the two backgrounds to give a common error of 0.0656 for 16 d.f. Applying these pooled errors, the one to the variance of sex means and the other to the variances of the joint means of the recombinant groups, brings out even more strongly the significance of the latter's increased variance: all comparisons are significant and five of the six at the 0.001 level.

One last test was made. The variances among parental means found in this experiment were lower than the variances among means of cultures of like genotype recorded earlier (table 2). The reason for this has not been established precisely, though several could be suggested. As a final test, therefore, the variances of the recombinant groups were compared with the higher values obtained earlier, viz. 0.4953 for males and 0.4173 for females. Since the variances are based on as many as 74 d.f., they can be divided into the corresponding sum of squares for the recombinant groups to give χ^2 s which show highly significant variation of both sexes on the B background. Of the two sexes, however, only the males show significance on the S background and that at no more than the 5 per cent. level. There is thus clear evidence of the effect of recombination in releasing variation in the X chromosome.

The third chromosomes depart from the X's in that the B and S lines differ very much in their chaeta producing powers in these chromosomes, whereas the X chromosomes for B and S were almost alike. The X's therefore provide a different and more sensitive test of the effects of recombination than do the chromosomes III. Indeed with a large initial difference between the B and S chromosomes III it would be possible for recombination in fact to reduce the overall variation by resolving coupling linkages.

The recombination experiment with chromosome III followed the same pattern as that with the X, the *Mé Sb* chromosome being used in extracting and fixing the prospectively recombinant chromosomes from the B/S heterozygote. Again 20 lines were used in each of the recombinant groups, the one on the B background and the other on the S. The parental group on each background included 5 lines with B and 5 with S chromosome III. The results were treated in the same way as those for the X and are similarly set out in table 10*a* and *b*.

From table 10*a* it will be seen that the overall means of the recombinant groups approximate well to the overall mean of the parental

groups on the S as well as on the B background. Turning to the variances of the means in table 10*b*, it is surprising to find these so low from the recombinant group on the B background. Those on the S background are much higher, but even so, they are less than the corresponding variances from the parent group. These latter are, of course, very much higher than in the corresponding X chromosome parental groups (table 9*b*) as must indeed be the case since they

TABLE 10

Results of chromosome III recombination

(a) Mean chaeta number with parental and recombinant X chromosomes

Background	Sex	Recombinant	Parents		Mean parent
			B	S	
B	Males . .	20.96	22.10	20.18	21.14
	Females . .	21.06	22.34	19.84	21.09
S	Males . .	19.26	20.58	18.14	19.36
	Females . .	19.69	21.16	18.18	19.67

(b) Variation among parental and recombinant cultures

Background	Sex	Mean squares	
		Recombinants 19 d.f.	Parental 9 d.f.
B	Males . . .	0.347	1.298
	Females . . .	0.612	2.032
	Joint . . .	0.306	1.579
S	Males . . .	1.695	2.027
	Females . . .	1.347	2.642
	Joint . . .	1.297	2.227

reflect the initial difference between the B and S chromosomes III. The indications are therefore that recombination is, if anything, reducing variation in respect of chromosome III, not releasing it as in the X. The effects of the two chromosomes on the D component will therefore be opposed, but the variance in the X would seem to be large enough to balance the decrease in the chromosome III, at least if we take the reduction observed on the S background as a fair measure. However this may be, it is clear that recombination is having its effects and that we cannot expect the D obtained by analysis, which

will reflect these effects, exactly to match the D synthesised from the consideration of whole chromosomes, which will not reflect the consequences of recombination. Fuller tests of the consequences of recombination, including consideration of chromosome II would be necessary to take further the exploration of the difference between the D values obtained analytically and synthetically, and to examine the effects on H .

5. THE NUMBER OF EFFECTIVE FACTORS

The number of effective factors, that is of effective units of segregation, may be estimated from biometrical data in a variety of ways (Mather, 1949). One of these is to divide the square of half the difference between the means of two true-breeding parental lines by the value found for D . If the segregating genes are isodirectionally distributed between the parents, that is all the $+$ alleles concentrated in one parent and all the $-$ alleles in the other, and there are genes all of equal d , half the parental difference is $S(d) = kd$ and $D = S(d^2) = kd^2$ giving

$$\frac{S^2(d)}{S(d^2)} = \frac{k^2 d^2}{kd^2} = k.$$

In so far as the distribution of the genes is not isodirectional, $S(d)$ and hence k are underestimated. Furthermore if the d 's vary from gene to gene k is also underestimated for when we set $d_a = \bar{d}(1 + \alpha_a)$; $d_b = \bar{d}(1 + \alpha_b)$; etc., so that $S(\alpha) = 0$ we find as the estimate of k

$$K_1 = \frac{k}{1 + V_\alpha} \text{ where } V = \frac{1}{k} S(\alpha^2) \quad (\text{Mather, 1949}).$$

The difficulty introduced by the failure of isodirectional distribution of the genes between the two parents is removed if the number of effective factors is estimated by one of a number of alternative approaches that depend on relations between the square of the mean variance and the variance of the variances in F_3 and second backcross families, for these depend on d^2 so that the sign of d ceases to influence the estimate. The estimate arrived at in this way, however, is overflow by an amount depending on V_β where, in the case of F_3 females $(d_a^2 + \frac{1}{2}h_a^2) = (\bar{d}^2 + \frac{1}{2}\bar{h}^2)(1 + \beta_a)$, $(db^2 + \frac{1}{2}h^2) = (\bar{d}^2 + \frac{1}{2}\bar{h}^2)(1 + \beta_b)$ etc. so that $V\beta$ must be greater than V_α and the underestimation correspondingly greater (Mather, 1949).

In our present data, we do not have F_3 families but BIP's so that the method applicable to the estimation of k for F_3 's cannot be used. An alternative, and in principle similar, method has been suggested by our colleague Mr R. Morley Jones.

As will be seen from table 11 the contribution of any gene to the variance of the variances of the BIP families (V_{vs}) can easily be shown

to be $\frac{1}{256} (8d^4 + 40d^2h^2 + 3h^4)$, so that summing over k genes of like d and h we find for the heritable portion of V_{vs}

$${}_H V_{vs} = \frac{k}{256} (8d^4 + 40d^2h^2 + 3h^4).$$

Now since $D = S(d^2) = kd^2$ and $H = S(h^2) = kh^2$, assuming all genes to have like d and h ,

$$\frac{1}{256} (8D^2 + 40DH + 3H^2) = \frac{k^2}{256} (8d^4 + 40d^2h^2 + 3h^4)$$

so that
$$K_s = k = \frac{8D^2 + 40DH + 3H^2}{256{}_H V_{vs}}$$

if all genes are of like effect. K_s will of course be an underestimate of k if there is any inequality of effect among the genes and this underestimation will depend on a quantity similar to V_β in that it depends

TABLE 11

The contribution of A—a to the variance of the variances of BIP families

Family	Frequency	V	V - $\bar{V}$	$(V - \bar{V})^2$
AA × AA	1	0	$-\frac{1}{4}d^2 - \frac{3}{16}h^2$	$\frac{1}{16}d^4 + \frac{3}{32}d^2h^2 + \frac{9}{256}h^4$
AA × Aa	4	$\frac{1}{4}(d-h)^2$	$-\frac{1}{2}dh + \frac{1}{16}h^2$	$\frac{1}{4}d^2h^2 - \frac{1}{16}dh^3 + \frac{1}{256}h^4$
AA × aa	2	0	$-\frac{1}{4}d^2 - \frac{3}{16}h^2$	$\frac{1}{16}d^4 + \frac{3}{32}d^2h^2 + \frac{9}{256}h^4$
Aa × Aa	4	$\frac{1}{4}d^2 + \frac{1}{4}h^2$	$\frac{1}{4}d^2 + \frac{1}{16}h^2$	$\frac{1}{16}d^4 + \frac{3}{32}d^2h^2 + \frac{1}{256}h^4$
Aa × aa	4	$\frac{1}{4}(d+h)^2$	$\frac{1}{2}dh + \frac{1}{16}h^2$	$\frac{1}{4}d^2h^2 + \frac{1}{16}dh^3 + \frac{1}{256}h^4$
aa × aa	1	0	$-\frac{1}{4}d^2 - \frac{3}{16}h^2$	$\frac{1}{16}d^4 + \frac{3}{32}d^2h^2 + \frac{9}{256}h^4$
	16	$\bar{V} = \frac{1}{4}d^2 + \frac{3}{16}h^2$	0	$\frac{1}{32}d^4 + \frac{5}{32}d^2h^2 + \frac{3}{256}h^4$

on variation among the d^2 s and h^2 s rather than among the d 's as does V_α .

Turning now to the estimation of the number of effective factors from the present data and taking K_1 first we note that the difference in chaeta number between the parental lines was 4.2690 for females and 4.4103 for males (table 2 of Cooke *et al.*, 1962). The weighted and unweighted estimates for D from the two sexes are given in table 6. Using the unweighted estimates and omitting the BS females for reasons already given above we then find

$$K_1 = \frac{2 \cdot 1345^2}{2 \cdot 96} = 1.54 \text{ for females}$$

and
$$K_1 = \frac{2 \cdot 2052^2}{2 \cdot 69} = 1.81 \text{ for males,}$$

and with the weighted estimates of D

$$K_1 = \frac{2 \cdot 1345^2}{2 \cdot 75} = 1 \cdot 66 \text{ for females}$$

and
$$K_1 = \frac{2 \cdot 2052^2}{2 \cdot 39} = 2 \cdot 03 \text{ for males.}$$

The evaluation of K_s requires a little more explanation in respect of ${}_H V_{VS}$. V_{VS} the variance of BIP variances is easily found from the experimental data and in, for example, the case of SB females it turns out to be 3.7804. This is, however, inflated by the sampling variance of the mean variance of the BIP families, V_{283} , which for SB females is 3.1315. This can be corrected by the subtraction of $\frac{2}{n-1} V_{283}^2$ where n is the number of flies from which the variance of any BIP family is calculated, in this case 10. The corrected value for SB females thus becomes $V_{VS} = 3 \cdot 7804 - \frac{2}{9} 3 \cdot 1315^2 = 3 \cdot 7804 - 2 \cdot 1792 = 1 \cdot 6012$. But even this value is too high for our purpose as we need ${}_H V_{VS}$ the *genetic* portion of V_{VS} . We can estimate the contribution which non-heritable agencies make to V_{VS} from the variance of the variances of parents and F_1 's, which must of course spring from non-heritable effects. Now the parents and F_1 's do not all show the same variances, so that E_1 , the non-heritable component of variation in the segregating families, has been taken as the average of the variances of the two parents and the two reciprocal F_1 's. In the case of SB females this was 2.3354 when averaged over all the parental and F_1 families raised. As Cooke *et al.* (1962) have recorded, however, single cultures of the two parents and two F_1 's were raised simultaneously whenever BIP families were being produced. We have therefore a series of observations from which the variance of this compound non-heritable variance can be found and it turns out to be 1.8110, for which an item, representing the sampling variance of the mean variance E_1 , must be subtracted in parallel to the correction applied above to V_{VS} . Since these variances were each based on 10 flies we find for the non-genetic component of

$$V_{VS} = 1 \cdot 8110 - \frac{2}{9} 2 \cdot 3354^2 = 1 \cdot 8110 - 1 \cdot 2120 = 0 \cdot 5990.$$

Subtracting this from V_{VS} we thus have for SB females

$${}_H V_{VS} = 1 \cdot 6012 - 0 \cdot 5990 = 1 \cdot 0022.$$

Now reference to tables 7 and 9 of Cooke *et al.* (1962) (where the results are given more extensively than in our table 6) shows that for these females $D = 2 \cdot 9550 \pm 0 \cdot 2387$ and $H = 3 \cdot 0118 \pm 0 \cdot 5902$ from the unweighted analysis and $D = 2 \cdot 7499 \pm 0 \cdot 4573$ and $H = 2 \cdot 5477 \pm 0 \cdot 9983$

for the weighted analysis. Then from the unweighted figures $8D^2 + 40DH + 3H^2 = 453.06$ and for the weighted figures

$$8D^2 + 40DH + 3H^2 = 360.20.$$

We can thus calculate

$$K_s = \frac{453.06}{256 \times 1.0022} = 1.77 \text{ from the unweighted estimates}$$

$$\text{and } K_s = \frac{360.20}{256 \times 1.0022} = 1.40 \text{ from the weighted estimates.}$$

Similar calculations can be carried out for the males and we find for SB males $K_s = 2.17$ using unweighted and $K_s = 1.70$ using weighted estimates and for BS males $K_s = 1.69$ using unweighted and $K_s = 0.34$ using weighted estimates. This last very low figure stems from the very low weighted estimate of $H(0.0944)$ from the BS males and it serves to emphasise that the lack of precision with which H , especially, is estimated from the data that are yielded by an experiment such as the present one must be reflected in the lack of precision of any quantity like K_s which depends on these estimates. The SB and BS males were separated in the estimation of K_s in order to bring out this point.

There are three major chromosomes in *Drosophila melanogaster* and all differing between the B and S stocks in their determination of sternopleural chaetæ albeit the effect of the difference attaching to the X chromosome is small. We should therefore expect k_1 the number of factors to be at least 3, or, if we regard the contribution of the X chromosome to the variation as negligible, at least 2. Equally, however, we know the effects of the chromosomes to differ in magnitude so that we expect both K_1 and K_3 to be underestimates. By how much should we expect them to underestimate K ?

Considering K_1 first, the direction of the X chromosome has been shown to be opposite to those of chromosomes II and III, in that the X chromosome for S is + relative to that from B, whose chromosomes II and III are + relative to those from S. K_1 will therefore be an underestimate because the genes are not distributed isodirectionally between the parents. The degree of underestimation is, however, hard to assess. If d_1 is taken as having the average of the male and female values of table 3, it will be only -0.030 whereas d_2 and d_3 are 1.741 , and the underestimation will be very small. If, however, we take for d_1 the value of -0.269 , as reached when the other data are included, the estimate of K_1 could be nearly halved from this cause. In finding the value expected for K_1 we have, however, assumed underestimation from this source to be negligible.

To find the extent of the underestimation springing from inequality of the effects of the factors we have assumed that each chromosome is acting as a unit, and that their relative effects are fairly represented

by the values for d_1 , d_2 and d_3 in table 3 averaged over the sexes. Both assumptions are over-simple. The d values of the table do not account for all the fixable heritable variation and if recombination is effective in releasing variation, so that the chromosomes are not acting as units and $k > 3$, and the values of d relative to one another cannot truly represent the inequalities.

In the absence of fuller information we proceeded by setting

$$d_1 = 0.0297 = \bar{d}(1 + \alpha_1); \quad d_2 = 0.6031 = \bar{d}(1 + \alpha_2); \\ d_3 = 1.1383 = \bar{d}(1 + \alpha_3)$$

giving $\alpha_1 = -0.9497; \alpha_2 = 0.0213; \alpha_3 = 0.9285$

and $S(\alpha^2) = 1.7645; V_\alpha = \frac{S(\alpha^2)}{k} = \frac{1.7645}{3} = 0.5882$

so that $K_1 = \frac{k}{1 + V_\alpha} = \frac{3}{1.5882} = 1.89.$

If we use all the data relating to the X chromosome and take $d_1 = 0.2694$ we similarly find $K_1 = 2.33$. These values are shown in columns A and B under "Expected" in table 12, alongside the average values

TABLE 12
The number of effective factors

Estimate	Observed			Expected with $k = 3$	
	Unweighted	Weighted	Pooled	A	B
K_1	1.68	1.85	1.76	1.89	2.33
K_S	1.88	1.15	1.51	1.52	1.65

Expected values in column A were calculated taking $d_1 = 0.030$.

Expected values in column B were calculated taking $d_1 = 0.269$.

observed for K_1 using the unweighted and weighted estimates of D , the column headed "Pooled" showing the average of the weighted and unweighted figures.

Turning to K_S , there is of course no underestimation arising from failure of isodirectional distribution of genes between the parents: we have only to consider the inequalities of effect among the factors. Here we encounter all the complications considered above in relation to K_1 , and in addition inequality of the h values also now comes in since H is explicitly involved in calculating K_S to whose value its contribution is by no means negligible, whereas it was not concerned in K_1 . Besides emphasising the possibility of recombination, the low values for the h 's in table 3 provide no good ground for assessing the relative contributions of the chromosomes to the dominance variation.

So the β values necessary for finding the extent to which K_s is made an underestimate by the inequality of effects of the factors have therefore been calculated from the d 's. This is obviously not strictly valid but it is the only course open in the absence of more reliable information about the values of the h 's relative to one another. Taking therefore the average d values from table 3 and recalling that we are now concerned with the relative values of the d^2 s rather than the d 's themselves we have

$$d_1^2 = 0.0009 = \bar{d}^2(1 + \beta_1); \quad d_2^2 = 0.3637 = \bar{d}^2(1 + \beta_2); \\ d_3^2 = 1.2957 = \bar{d}^2(1 + \beta_3)$$

giving $\beta_1 = 0.9998; \beta_2 = -0.3428; \beta_3 = 1.3413$

$$S(\beta^2) = 2.9162 \quad V = \frac{S(\beta^2)}{k} = \frac{2.9162}{3} = 0.9721$$

so that $K_s = \frac{k}{1 + V_\beta} = \frac{3}{1.9721} = 1.5212.$

If we take $d_1 = 0.2694$ we similarly find $K_s = 1.6513$. These also are entered in columns A and B respectively under "Expected" in table 12, alongside the average values observed for K_s set out just as for K_1 .

The figures in table 12 require little comment: the agreement between the values expected and observed is surprisingly good when we bear in mind the high error variance to which the observed values must be subject and the over-simple assumptions on which the expected figures are based. The values observed for K_1 are somewhat below expectation which presumably reflects in part the lack of isodirectional distribution between the parents and partly the overlow value of 3 assumed for k : recombination must result in k exceeding the number of major chromosomes. The shortfall in the observed K_s is less striking, suggesting that perhaps the assumption of $k = 3$ is not grossly misleading and that the difference in K_1 is due more to the lack of isodirectional distribution between the parents. The overall impression is, however, that the values calculated for the K 's from the observed components of variation are reasonable. This is particularly encouraging in the case of K_s . Estimates of the number of effective factors obtained from the variances of segregating generations have often in the past been absurdly low: we may now feel some confidence that with proper safeguards and corrections such calculations can yield reasonable results.

6. SUMMARY

Treating the three major chromosomes as units of inheritance, the eight possible true-breeding lines have been built up using the chromosomes of the B and S stocks of *Drosophila melanogaster*. These eight lines have been mated in all possible combinations to give the

27 genotypes possible in females and 18 in males. The numbers of sternopleural chaetæ associated with these genotypes have been counted and from them were estimated the 26 genetic parameters in females and the 17 in males which completely describe the effects of these chromosomes and their interaction on the variation in this character (table 3).

The values of $D = S(d^2)$ and $H = S(h^2)$ have been synthesised from the estimates of the parameters and compared with the values of D and H obtained by Cooke *et al.* (1962) from the biometrical analysis of chaeta number in B and S lines, their F_1 , F_2 and BIP generations. The synthetic values of D in the two sexes are only about 60 per cent. of the values found by biometrical analysis. The difference is even more striking in H , but the high standard errors of the analytical values render interpretation more hazardous. The difference in D may be due to the effects of recombination within the chromosome. Any variation released by such recombination would raise the value of D and H obtained from the biometrical analysis by comparison with the synthetic values which were derived by treating each chromosome as an indivisible unit.

The number of effective factors was estimated both from the parental differences (as K_1) and from the variances of the BIP progenies (as K_s) in Cooke *et al.* (1962) experiment. These estimates were compared with expectations based on an assumed actual value of 3 (taking each chromosome as a unit) but with allowance made for underestimation arising from differences in the effects of the factors as exemplified by differences among the d values estimated for the major chromosomes. Agreement between observation and expectation was good, so encouraging confidence in the methods used for estimating K_1 and especially K_s biometrically.

7. REFERENCES

- COOKE, P., JONES, R. M., MATHER, K., BONSALE, G. W., AND NELDER, J. A. 1962. Estimating the components of continuous variation. I. Statistical. *Heredity*, 17, 115-133.
- FISHER, R. A. 1953. *The Design of Experiments*. 6th ed. Oliver and Boyd, Edinburgh.
- HAYMAN, B. I., AND MATHER, K. 1955. The description of genic interactions in continuous variation. *Biometrics*, 11, 69-82.
- MATHER, K. 1943. *Statistical Analysis in Biology*. Methuen, London.
- MATHER, K. 1949. *Biometrical Genetics*. Methuen, London.
- NELDER, J. A. 1960. The estimation of variance components in certain types of experiment on quantitative genetics. *Biometrical Genetics*, Pergamon Press. 139-158.