THE INHERITANCE OF KERNEL WEIGHT IN A *TRITICUM VULGARE* CROSS

L. G. L. COPP
Crop Research Division
and
G. M. WRIGHT
Crop Research Division and Applied Mathematics Laboratory, Department of Scientific and Industrial Research, Christchurch, New Zealand

1. INTRODUCTION

THE study of the inheritance of a complex character, such as yield in wheat, is made difficult by variation caused by differences in the environment, and by the effects of interaction of environment and genotype. Yield per plant may be resolved into a number of sub-characters, or components, which may not be independent in their genetic control (Mather, 1949). The major components, each capable of further resolution, are number of ears per plant, number of kernels per ear, and average weight of a single kernel. Kernel weight has been shown to be the least variable of these components (Smith, 1936).

Estimates of the numbers of genes which determine kernel weight differences in wheat have already been published. Jasnowski (1934, 1935), quoted by Boyce (1948), obtained evidence that certain lines differed by three pairs of genes. Worzella (1942) stated that the inheritance of kernel weight in a number of wheat (*Triticum aestivum* L.) hybrids was multigenic. Boyce (*l.c.*), in a preliminary study, by empirical methods, of the material described in the present paper, concluded that the F_2 and F_3 frequency distributions, and the difference in kernel weight between the two parents, were the result of the action of two or of three major gene pairs.

The plants were grown at the Wheat Research Institute, Lincoln, in 1943/44, but a detailed analysis was not then carried out. Interest in the material was revived following the publication in 1949 of Mather's *Biometrical Genetics*.

2. MATERIAL AND METHODS

The material was one of four crosses used in a preliminary study of the inheritance of kernel weight (Boyce, *l.c.*), comprising the parents, F_1, F_2 and reciprocal F_3 generations of the *Triticum vulgare* cross (India 29, S1556 × Crete 2, S357). Where P_1 (S1556) was used as pollen parent the hybrid material was designated as F_2A or F_3A. The parents were chosen for their high and low kernel weights from a collection of foreign varieties. The purity of the lines had been maintained for several years by sowing grain from ears which had been covered before

187

anthesis.   Each F_3 plot of 52 plants was the progeny of a single F_2 plant.   The positions of eight F_1 plots of 26 plants and eight F_2A plots of 52 plants were randomized among the F_3 plots, which were not replicated.   Parent plots of 26 plants were grown in random order after every second hybrid plot.   Evenness of spacing and depth were obtained by sowing individual grains with a " Woodfield " dibber. The grains were sown two inches deep, with eight inches between the rows of thirteen plants four inches apart.

Plants were harvested individually.   For each plant the number of ears (e), total number of kernels (en), and the yield (eng) were recorded; the average number of kernels per ear (n) and the average weight of one kernel (g) were calculated.   Kernel weights for each plant were expressed in milligrams, and ranged from 20 to 61.

Plants with only one ear, or with fewer than forty kernels (about five per cent of those harvested in each population), were omitted from the analysis because average kernel weight would not have been determined with the same accuracy for these plants as for the rest of the material, and because, in many of them, the grain was shrivelled through disease or poor development.   The analysis, therefore, is based on healthy plants only.

In studies of quantitative inheritance the amount of material is usually limited, and to obtain the maximum information (from the segregating populations) relatively small parent and F_1 populations are grown.   The size of the F_1 is also limited by the work involved in crossing.   For any character, all population frequency distributions are commonly presented in one figure.   If the same frequency scale were used for all populations it would be difficult to make visual comparisons between them because the differences between the sizes of the popu-
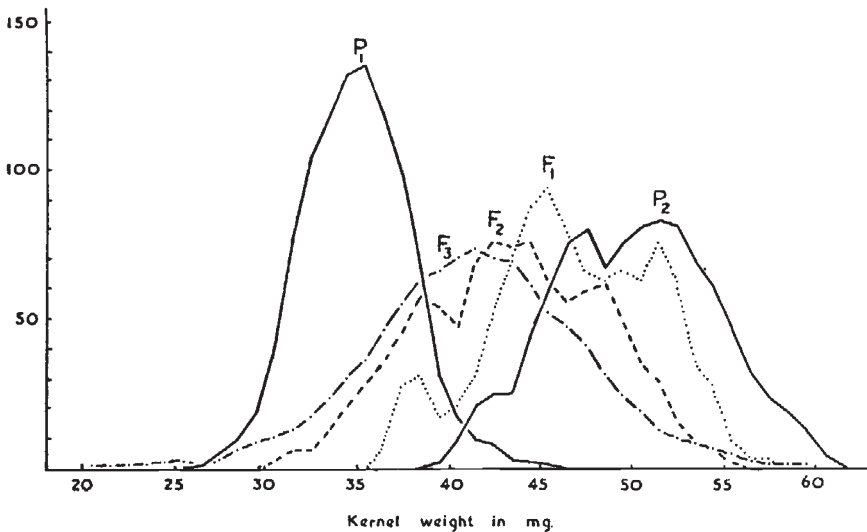

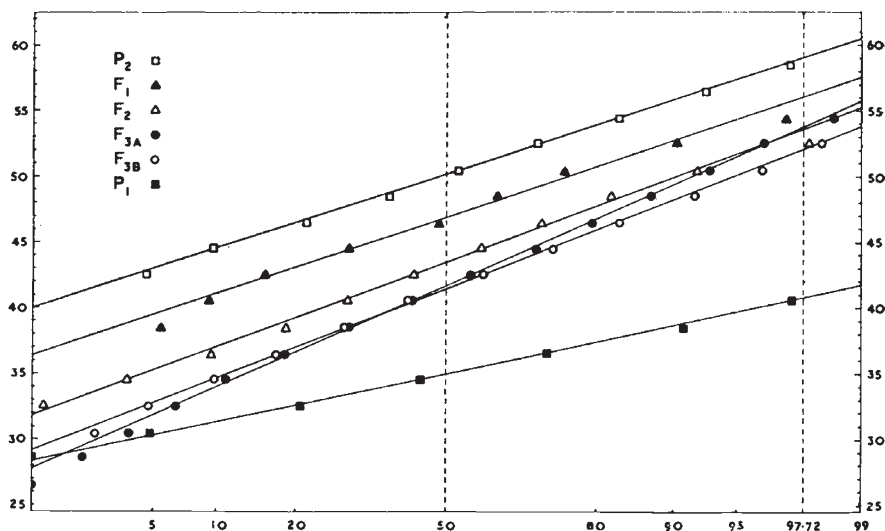Fig. 1.—Frequency distributions adjusted to equal population size.

Fig. 2.—Cumulative frequency distributions of kernel weight on probability scale. The mean is located on the ordinate at 50%. The standard deviation is proportional to the slope of the line.

lations are large. A second method, still complicated, in small populations, by the relation between range and population size, would be to adjust the populations to equal total frequencies. In figure 1 a two-term moving average of adjusted kernel weight frequencies has been plotted.

A limitation to these methods of presenting the data is the difficulty experienced in distinguishing and comparing the population means and variances, but on a cumulative probability scale, by which normal distributions are represented by straight lines, the means and variances of the populations are easily compared. In figure 2, which was drawn on probability graph paper, there is, for each population, a point for every second unit of kernel weight over the population range. The numbers of plants in each kernel weight group were accumulated from the lower end of the range and the figures so obtained were expressed as percentages of the total number of plants. These percentages were plotted on the probability scale against the kernel weight, at the midpoints of the class intervals. For example, the percentage with kernel weight 44 or less (which means, in fact, less than 44·50) was plotted at 44·5: if it were plotted at 44·0 and the other points were plotted similarly, the mean, read off at the 50 per cent point, would be one half of the unit of measurement, or 0·5 mg., below the calculated mean. This method of presentation was used by Keller and Li (1949), who did not make this adjustment.

The normal distributions fitting each population are depicted in figure 2 as straight lines fixed by plotting the population mean at 50% probability, and the point twice the standard deviation above the mean at 97·72% probability.

3. RESULTS

(i) *Description of Populations*

The populations had distributions close to normal (figure 2). The six population means and variances and some correlations of the raw data are given in table 1*a*. The parent (S357) with the higher kernel

TABLE 1

Population Statistics

(a) *Average kernel weight (g); mg. per grain*

Generation . .	P_1	P_2	F_1	F_2A	F_3A	F_3B
Mean per plant .	34·96	50·10	46·88	43·41	41·59	41·40
Standard error of mean	0·11	0·13	0·37	0·31	0·106	0·097
Population standard deviation . .	2·86	4·36	4·54	5·02	6·025	5·281
Variance between plots .	*41·48***	*131·98***	*34·80*	*71·29***	*481·90***	*283·48***
Variance within plots .	*5·80*	*12·48*	*19·91*	*23·78*	*22·04*	*20·06*
Parent-offspring correlation .	0·304	− 0·264	—	—	0·737**	0·438**
Number of plots .	34	40	8	8	62	64
Number of plants .	497	711	143	242	1968	2120

(b) *Transformed variate: log$^{\frac{3}{2}}$ (g − 25)*

Mean per plant .	0·9808	2·3022	2·0591	1·7711	1·6021	1·5925
Standard error of mean	0·0117	0·0098	0·0295	0·0270	0·0095	0·0089
Population standard deviation . .	0·3095	0·3157	0·3575	0·4317	0·5416	0·4791
Variance between plots .	*0·4839***	*0·6412***	*0·1995*	*0·5034***	*3·8999***	*2·2943***
Variance within plots .	*0·0681*	*0·0682*	*0·1241*	*0·1769*	*0·1779*	*0·1662*
Parent-offspring correlation .	0·251	− 0·280	—	—	0·727**	0·407**

weight had a greater variance within plots than the other parent $(F = 2·15**)$, but there was no correlation between plot means and variances in any population. Accordingly, the linear scale of measurement would be considered satisfactory for descriptive purposes. However, for a biometrical analysis of the data, the scale should be such that gene effects appear additive (Mather, 1949, ch. 3) and such that there is no apparent interaction of environment and genotype. The fulfilment of these conditions depends on the positions of the F_2 and F_3 means in relation to those of the parents and F_1, and on the comparison of the variances of the three genetically homogeneous populations, P_1, P_2, and F_1.

From table 1*a*, Mather's $C = 4\bar{F}_2 - 2\bar{F}_1 - \bar{P}_1 - \bar{P}_2 = -5·3$

and $D = 8\bar{F}_3 - 2\bar{F}_1 - 3\bar{P}_1 - 3\bar{P}_2 = -16·3$

and using variances within plots $V_C = 2·16$ and $V_D = 1·48$, so that C and D (hence $\bar{F}_2$ and $\bar{F}_3$) were significantly lower than would be expected if gene effects were simply additive. In fact, $\bar{F}_3$ was slightly, and significantly, lower than the mid-parent.

(ii) *Scaling*

It may be seen also that the within plot variances of P_1, P_2 and F_1 were unequal, so that neither criterion for the scale was satisfied. In the F_1, more than the scale was at fault, for the plot variances were not homogeneous. A suitable foreshortening scale would equalise the parent variances and improve the relative positions of the means, although it would affect the normality of the distributions. A square root scale was not sufficiently powerful. It was decided to use a modified logarithmic transformation deduced as follows: writing $\overline{G} = \log_{10}(\bar{g} - a)$ for $\bar{g}$ in the expressions for C and D above, where g is the original kernel weight, G the weight on the scale to be found, a a constant to be determined, it was found that C changed sign for a between 25 and 30, and D for a between 20 and 25, so that $\log(g - 25)$ seemed a satisfactory variate for analysis. The use of $\log(\bar{g} - a)$ rather than the mean of $\log(g - a)$ in calculating C and D made the approximate determination of a feasible. Unfortunately, on this scale the P_1 variance was higher than that of P_2 ($F = 3 \cdot 10^{**}$), and as a compromise a less powerful transformation was needed. Use of $\log(g - 10)$ equalised the parent variances but was not satisfactory for the means, $\log^2(g - 25)$ gave a smaller variance ratio ($F = 1 \cdot 38^{**}$) for the parents than did $\log(g - 25)$, and by extrapolation $\log^{\frac{3}{2}}(g - 25)$ was thought likely
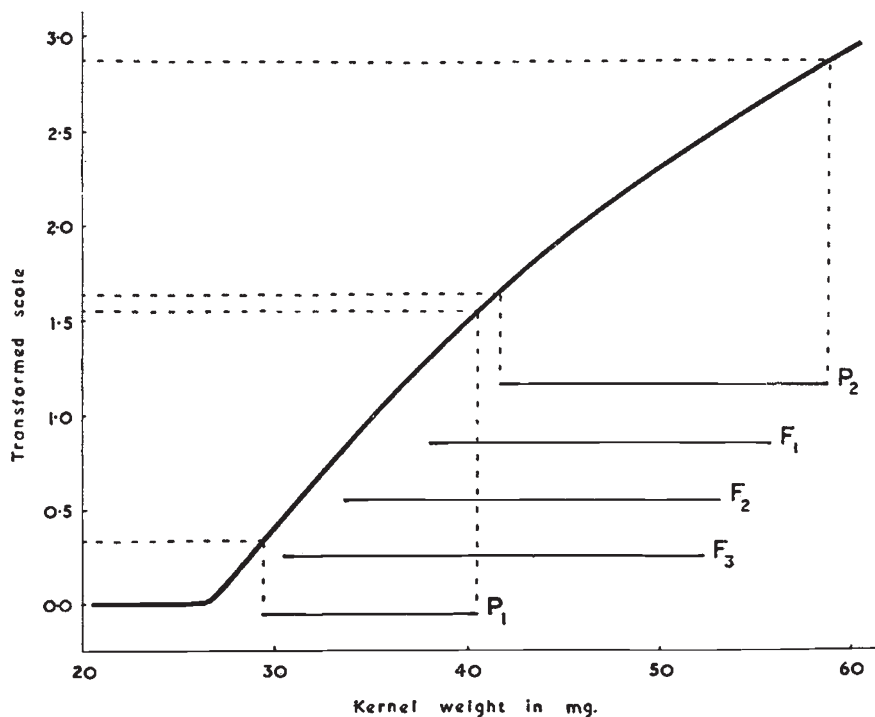

FIG. 3.—Graph of the function $y = \log^{\frac{3}{2}}(g - 25)$ for transforming kernel weight figures. Ranges covering 95 per cent of the population distributions. P_1 and P_2 ranges are shown on both scales.

to be adequate. This function is plotted in figure 3. The ratio of parent variances with this transformation was 1·04.

Final calculations of C and D gave $-0·317$ and $-1·138$, with variances respectively 3·46 and 12·82, which is satisfactory though the F_3 mean is still slightly below the mid-parent. Once it has been decided to use a particular, arbitrary, transformation, its theoretical form is of little interest. It is unlikely that the same transformation would have been appropriate in another year, as in both the preceding and following years the average kernel weights of the parent varieties, grown in the same locality, were significantly higher.

(iii) Transformed Data

The statistics derived from the transformed data are shown in table 1b.

As already stated, the F_2 and F_3 means are now closer to their expected positions in relation to the means of the parents and F_1. The within-plot variances of the parents have been equalised. The F_1 variances, based on only 143 plants, have not been used in the biometrical analysis because they are still significantly different from those of the parents. The transformation has had no appreciable effect on variance ratios or on the parent-offspring correlations. Except in F_1, the between-plot variances are highly significantly greater than those within plots. The large variance between plots of the parents is attributed to soil heterogeneity.

The F_2 variance between plots is similar to those of the parents, the only genetic differences between F_2 plots being due to sampling. The within-plot variances of the F_2 and F_3, which express genetic variation, are much larger than those of the parents.

Although there is no significant difference in average kernel weight between the reciprocal F_3 populations, their statistics have been presented separately. There are differences in the parent-offspring correlations and in the between-plot variances: the former difference is significant, and the latter almost significant, at 5 per cent. A highly significant difference between the reciprocal F_3's in average number of kernels per ear is not discussed in the present paper. In the previous year six F_2 plots of approximately 60 plants had been grown, three F_2A and three F_2B. A random selection of these plants was used for the F_3; their average kernel weights for A and B were similar to those of the remaining plants. The over-all variance of the F_2A, as of the sample grown, was considerably greater than that of the F_2B ($F = 1·37$, significant at 5 per cent). This difference, for which there is no explanation, would have contributed to the difference in F_3.

(iv) Biometrical Analysis
(a) Components of Variation

The second-degree statistics of segregating populations are functions of D, H and E, which are components of variation dependent on the

fixable and the unfixable genetic variation and the environmental variation respectively. The required statistics have the following expression (Mather, *l.c.*):

$$V_{F2} = \tfrac{1}{2}D + \tfrac{1}{4}H + E_1, \text{ the within plot variance of } F_2 \text{ plants,}$$

$$V_{\overline{F3}} = \tfrac{1}{2}D + \tfrac{1}{16}H + E_2, \text{ the variance of } F_3 \text{ progeny means,}$$

$$W_{F2/F3} = \tfrac{1}{2}D + \tfrac{1}{8}H, \text{ the covariance of } F_2 \text{ plants of the previous year}$$
$$\text{with their progeny means,}$$

$$\overline{V}_{F3} = \tfrac{1}{4}D + \tfrac{1}{8}H + E_1, \text{ the mean variance of the } F_3 \text{ progenies.}$$

E_1 and E_2 are the non-heritable components of variation for single plants and for means of plots, estimated directly from the non-segregating populations. The least squares calculation of D, H, E_1 and E_2 from these six relations is described by Mather (p. 65). The observed "A" and " B " second-degree statistics and their averages, and the calculated values of D, H, E_1 and E_2 are presented in table 2.

TABLE 2

Observed second degree statistics, derived estimates of genetic and environmental components, and number of effective factors

	A	B	Combined
Statistics			
V_{F2}	0·176913	—	$\begin{cases} 0 \cdot 262372* \\ 0 \cdot 207054 \end{cases}$
$V_{\overline{F3}}$	0·131369	0·070628	*0·100999*
$W_{F2/F3}$	0·150489	0·051722	*0·101106*
V_{F3}	0·181914	0·166120	*0·174017*
E_1	0·068185	0·068139	*0·068162*
E_2	0·021414	0·016159	*0·018786*
V_{VF3}	0·436046	0·532727	*0·484387*
Components of variation			
D	0·280± ·129	−0·017± ·122	*0·063± ·059*
H	−0·072± ·413	0.627± ·489	*0·628± ·238*
E_1	0·081± ·035	0·080± ·027	*0·074± ·013*
E_2	0·009± ·035	0·028± ·027	*0·025± ·013*
$D+\tfrac{1}{2}H$	0·244± ·106	0·296± ·140	*0·377± ·068*
Number of effective factors			
K_1	1·56 (from D)	0·27 (from H)	*1·39 (from $D+\tfrac{1}{2}H$)*
K_2	0·009	0·012	*0·018*

* These two figures are estimated from D, H, and E_1, from the " combined " estimates and the " restricted " linkage estimates respectively. There was no " observed " figure.

For the separate "A" and " B " analyses the component E_1 was estimated directly as the within-plot variance of P_2 or P_1: E_2 was estimated as the between-plot variance of P_2 or P_1, divided by the harmonic mean (29·94) of the numbers of plants per F_3 plot. Since no F_2B was grown, for the " B " and " combined " calculations it was necessary to prepare and invert a new matrix of coefficients. This matrix and its inverse are given in table 3.

TABLE 3

Matrix used in " B " and " combined " analyses in calculating
D, H, E_1 and E_2 without V_{F2}

$$M = \frac{1}{256} \begin{vmatrix} 144 & 32 & 64 & 128 \\ 32 & 9 & 32 & 16 \\ 64 & 32 & 512 & 0 \\ 128 & 16 & 0 & 512 \end{vmatrix} = \begin{vmatrix} 0\cdot5625 & 0\cdot125 & 0\cdot25 & 0\cdot5 \\ 0\cdot125 & 0\cdot03515625 & 0\cdot125 & 0\cdot0625 \\ 0\cdot25 & 0\cdot125 & 2\cdot0 & 0 \\ 0\cdot5 & 0\cdot0625 & 0 & 2\cdot0 \end{vmatrix}$$

$$M^{-1} = \frac{1}{25} \begin{vmatrix} 416 & -1536 & 44 & -56 \\ -1536 & 6656 & -224 & 176 \\ 44 & -224 & 21 & -4 \\ -56 & 176 & -4 & 21 \end{vmatrix} = \begin{vmatrix} 16\cdot64 & -61\cdot44 & 1\cdot76 & -2\cdot24 \\ -61\cdot44 & 266\cdot24 & -8\cdot96 & 7\cdot04 \\ 1\cdot76 & -8\cdot96 & 0\cdot84 & -0\cdot16 \\ -2\cdot24 & 7\cdot04 & -0\cdot16 & 0\cdot84 \end{vmatrix}$$

In table 2 it will be seen that the results from the "A" and " B " analyses are different, particularly with respect to the values of H. The " combined " analysis gave results very similar to the " B " analysis, but agreeing much more closely with the observed statistics. This similarity may be merely an effect of the omission of F_2 data. The components of variation were not accurately determined. (The "A" standard errors have two degrees of freedom, the others one.) There is therefore little justification for their use in calculating the number of effective factors which differentiate the two parents, or the degree of dominance or potence shown. Mather has suggested that, in view of the negative correlation between D and H as shown in the two halves of the experiment, an estimate of $D + \frac{1}{2}H$ should be more consistent. The figures given in table 2 support this, and it will be seen that $D + \frac{1}{2}H$ is estimated with the same precision as D. The variance of the estimate depends on c_{DD}, c_{DH} and c_{HH} of the inverse matrix.

(b) Test for Linkage

A modification of Mather's " restricted " analysis, in which the mean variance of F_3 plots is omitted, is available as a test for linkage, though in the present material it would not be sensitive. In the "B" or " combined " analyses there would remain only four statistics involving D, H, E_1 and E_2, and therefore no residual degrees of freedom. If the calculated " combined " values of D, H, E_1 and E_2 are used to estimate the second-degree statistics, as in calculating their standard errors, V_{F2} may be estimated also, and the eleven observed "A" and " B " statistics may be compared with these estimates. The sum of squares of the eleven deviations, with seven degrees of freedom, is the total sum of squares in the linkage analysis. Similarly a set of statistics may be estimated from the " restricted " values of D, H, E_1 and E_2, giving the remaining value of V_{F2} in table 2. Since the four variance components were calculated from only four statistics, the remaining estimates are identical with the observed figures. The sum of squares of the eleven differences between these figures and the observed "A" and " B " statistics is a residual sum of squares; the difference between the total and the residual is the sum of squares, with one degree of freedom, for linkage. The ratio of the linkage and residual *mean* squares was 5·27 (F.₀₅, 5·99) which suggests that there is linkage between the

genetic factors influencing kernel weight.   However, in the absence of an F_2B this difference in variance cannot be attributed definitely to linkage.

(c) Number of Effective Factors

The number of " effective factors " differentiating the parents may be estimated by the use of components of variation and the parent and F_1 means (Mather, *l.c.*, p. 102).   If k effective factors have equal effects d, all in one direction, the expected value of $(P_2 - MP)^2$ is k^2d^2, which equals kD, so that k may be estimated as $K_1 = \dfrac{(P_2 - MP)^2}{D}$ and similarly, if H is known accurately, $K_1 = \dfrac{(F_1 - MP)^2}{H}$, where P_2 and F_1 represent the corresponding averages and $2MP = P_1 + P_2$.   No precise estimate of K_1 can be obtained from this material; the best estimate is $\dfrac{2(P_2 - MP)^2 + (F_1 - MP)^2}{2D + H}$, which, using $(2D + H)$ from the " combined" analysis, equals $1 \cdot 39$.   If, as is likely, the parents differed in many effective factors, an estimate of this order for K_1 would indicate incomplete concentration, the parents having both plus and minus allelomorphs.   The data are compatible with a K_1 as large as twelve.

A second estimate, K_2, is derived from the variances of the F_3 plots, as the ratio of the square of the heritable part of the mean variance of the F_3 plots $(\tfrac{1}{4}D + \tfrac{1}{8}H)$ to a corrected variance of the plot variances $(V_{VF3}$, table 2; Mather, *l.c.*, p. 104).   This K_2 is expected to be smaller than K_1, except when K_1 is reduced by incomplete concentration, and the relative smallness of the K_2 as shown in table 2 would indicate, if the estimates were reliable and there were in fact many effective factors, that the factors have unequal effects on kernel weight.   But the number of effective factors by which the parents differ remains undetermined, and, in consequence, no reliable estimate can be obtained of the immediate limits, or speed, of selective advance.

(d) Difference between Reciprocal Crosses

The components of variation estimated from the two halves of the experiment give two different estimates of the genetic difference between the parents.   If the discrepancy is real (and the differences in components are certainly not significant) there is a difference between reciprocal crosses.   There is a real difference in two of the second-degree statistics, the greater range of the means of the F_3A plots partly explaining their higher covariance with their parent plants.   Similarly, the variances of the F_2A progenies in the previous year were greater than those of F_2B.

(e) Results in Other Years

The above analysis has been confined to the results obtained in 1943/44, when the difference in kernel weight between the two parents

was $15\cdot1 \pm 0\cdot1$. In the previous year, under much better growing conditions, the difference was $15\cdot9 \pm 0\cdot3$, and the higher parent averaged $57\cdot3 \pm 0\cdot2$ (compared with $50\cdot1 \pm 0\cdot1$). The relation of the F_1 and F_2 means to the parent means was similar in the two seasons. Three of the F_2 plants were selected for further study, two with average kernel weight slightly greater than P_2 and one much lower than P_1. Their F_3 progenies, grown as part of the main experiment, had average kernel weights much closer to the F_3 mean than the original plants were to their F_2 mean, although the third progeny was again lower than P_1. The progenies of four plants from each of these F_3 plots were grown in the following year, again under good conditions, the F_3 plants in each plot being selected to cover the plot range of kernel weight, and averaging close to the plot average. In this F_4 the kernel weights of the progenies derived from the two high F_2 plants were all nearly as high as the high parent, which averaged $51\cdot7 \pm 0\cdot2$, but the progenies from the low F_2 plant were significantly above the low parent. The original low F_2 plant was apparently strongly affected by its environment, yet the single-plant selection in F_2 was effective, and much more effective than selection in F_3. This would suggest that the number of effective factors by which the parents differed was not large.

4. DISCUSSION

In this experiment, as in some of those listed by Mather, the components of variation were not estimated precisely, although much material was grown. The design of the experiment did not satisfy all the requirements for biometrical analysis prescribed by Mather; the F_3 plots were not replicated, the F_2 and F_3 plots were twice as large as those of the F_1 and the parents, and the two parents were sown in random order at regular intervals. Only the first of these defects is serious, but it has been pointed out recently (Mather and Vines, unpublished) that more information will be obtained by the growing of a given number of unreplicated F_3 families than from half that number in two replications.

The plots were arranged in eight adjacent blocks, but the parent plots showed no significant block differences. Block differences in the F_3 would be confounded with genetic differences between plots. Replication is desirable, to supply a more precise estimate of the errors of the components of variation and for the test of linkage. Unfortunately, different numbers of plots of the P_1 and F_3 were harvested from each block, so that an analysis with eight " replicates " would not have been justified. The separation of the reciprocal crosses in the present material provided in the second-degree statistics a set of eleven deviations. The standard errors of the " combined " estimates of D, H, E_1 and E_2 derived from the deviations from the " combined " second-degree statistics had only one degree of freedom. If the deviations from the separate "A" and " B " statistics had been used, there would have been seven degrees of freedom available, but the differences be-

tween the reciprocal crosses were such that the test of significance of the components would not have been improved.

It was not possible to find a scale of measurement on which the parent variances would be equalised and the F_3 means would be higher than the mid-parent, though on the scale adopted the F_3 was not significantly below its expectation. The form of the transformation is restricted mathematically only by the condition that it must be continuous and monotonically increasing over the range of expression. In studies of quantitative inheritance there is no reason to expect that one of the " standard " transformations will be the most appropriate. A biometrical analysis of the untransformed data led to conclusions similar to those presented; the transformation made the data conform more closely to the theoretical genetic criteria but its use has not provided more, or more precise, information. It might be suggested that the F_1 mean should have been ignored in testing the scale, since the F_1 plot variances were not uniform, but the mean itself was accurately known, and the F_3 mean was still significantly below expectation based on F_2 and parent means. The most satisfactory effect of the re-scaling, forced to equalise the *within-plot* variances of the parents, was the equalising of the *between-plot* variances of the parents and F_2. Their ratios were changed from $1 : 3 \cdot 18 : 1 \cdot 72$ to $1 : 1 \cdot 33 : 1 \cdot 04$.

The difference between the reciprocal crosses affects the estimates of D and H, but not $D + \frac{1}{2}H$. It has been stated that the difference may be attributable to a difference in the F_2 of the previous year, but the origin of this difference is unknown. It is unfortunate that no F_2B was grown with the F_3's. The original F_2 difference might have been a consequence of differential incidence of seed-borne (and hence " maternally inherited ") disease in the first F_1 plants.

In this material the weighting of the second-degree statistics was not conducive to the improvement of the estimates. Weighting the observations E_1 and E_2 heavily did not increase the accuracy of estimation, which suggests that the re-scaling had not removed all interaction of genotype with environment. A disadvantage of weighting is that the required matrix of coefficients is not one of the standard forms, and is not symmetrical. The standard matrices are relatively easy to invert by Fisher's method, particularly if the equations set up are cleared of fractions; the terms of the inverse matrices are found as fractions with a common denominator which is, apart from powers of two, the value of the determinant of the original matrix. Thus in Mather's table 12 (p. 61) c_{HH} is exactly $\dfrac{8704}{93}$, which equals $93 \cdot 591398$, and not $93 \cdot 591441$ as given.

Kernel weight was shown by Smith (1936) to be the most effective yield component for differentiating between certain varieties of wheat. The present parents were chosen because they differed widely in average kernel weight: $98 \cdot 6$ per cent of the parent plants could be classified correctly on kernel weight alone, and about half of the remainder were

classified correctly if number of kernels and number of ears were also taken into account.

Boyce (1948) showed that the shape of the F_2 distribution of kernel weight could be explained by the segregation of two partially dominant genes, but she was unable to find any satisfactory explanation for the F_3 distribution. No doubt it would be impossible to get a statistically satisfactory fit in this large F_3 (Haldane, 1949). Formal estimates of number of genes, or effective factors, are unreliable, but they are valuable in indicating whether the effects of genes are equal and in the same direction. The data suggest that the effective factors influencing kernel weight are possibly linked, have unequal effects, and do not all act in the same direction.

Effectiveness of *selection* for yield components may be less than for total yield (Boyce *et al.*, 1947). In the material of Palmer (1952), although selection for kernel weight was effective, it had no effect on total yield, because kernel weight was negatively correlated with number of kernels per plant. The present F_4 data show that a line with kernel weight as high as the high parent is recoverable, and that it may be possible to select a line as low as the low parent. This information about the limits of selection is more definite than that obtained from the biometrical analysis.

The application of Mather's methods of analysis has not been successful. The variances and covariances of the segregating populations did not have the general relationships expected on simple genetic reasoning: the F_2/F_3 covariance was relatively high in the "A" material and low in the " B ". There are several possible explanations. The difference between parental genotypes is fixed, but would have a different expression in another environment. Some mistakes would have been made in counting, weighing and computing, and these would have increased the error variances even if no bias were introduced. Non-heritable variation, sampling variation and residual genic interaction would have lowered the precision of the analysis. Sampling variation would be expected to be small in this large material, and non-heritable variation appears to have been adjusted by the re-scaling, leaving residual genic interaction as the most probable explanation for the lack of success. There is no suggestion that any other method of analysis would have given more information.

5. SUMMARY

1. The inheritance of kernel weight was studied in the first three hybrid generations of a varietal cross in *Triticum vulgare*, using the biometrical methods developed by Mather.

2. A transformation was derived by which the data satisfied Mather's two criteria for an adequate scale of measurement. The empirical method of obtaining the transformation, by means of the formulae for the tests of a scale, has been explained.

3. The advantages of presenting a set of frequency distributions on probability graph paper have been discussed.

4. Although the material was large the estimates of the genetic and environmental components of variation were not sufficiently precise to enable any definite conclusions to be drawn regarding the number of effective factors segregating, their linkage relationships or the possible limits to selection. It has been suggested that residual genic interaction has affected the precision of the estimates.

5. There were unexplained differences between the distributions of the reciprocal F_3 progeny means, which produced differences between the separate estimates of the components of variation.

REFERENCES

BOYCE, S. W. 1948. A preliminary study of the inheritance of grain weight in wheat. *N.Z.J.Sci. and Tech., 30*, 13-22.

BOYCE, S. W., COPP, L. G. L., AND FRANKEL, O. H. 1947. The effects of selection for yield in wheat. *Heredity, 1*, 223-234.

HALDANE, J. B. S. 1949. Some statistical problems arising in genetics. *J. Roy. Stat. Soc. B, XI*, 1-14.

JASNOWSKI, S. 1934. On the inheritance of weight in grains of *T. vulgare* L. *Roczn. Nauk. Rolnicz. i Lesnyck, 33*, 59-68.

JASNOWSKI, S. 1935. On the inheritance of some characters of the ear of wheat. *Polon. Acad. Bul. Sc. Lettres Carcovi*, 199-218.

KELLER, K. R., AND LI, J. C. R. 1949. The relationship between the number of vines per hill and yield in hops. *Agron. J., 41*, 569-573.

MATHER, K. 1949. *Biometrical Genetics*. London.

MATHER, K., AND VINES, A. Inheritance of height and flowering time in a cross of *Nicotiana rustica*. *Heredity* (in the press).

PALMER, T. P. 1952. Population and selection studies in a *Triticum* cross. *Heredity*, (in the press).

SMITH, H. FAIRFIELD. 1936. A discriminant function for plant selection. *Ann. Eugen., 8*, 240-250.

WORZELLA, W. W. 1942. Inheritance and inter-relationship of components of quality, cold resistance and morphological characters in wheat hybrids. *J. Agr. Res., 65*, 501-522.