RELEASE OF GENETIC VARIABILITY THROUGH RECOMBINATION.

IV. STATISTICAL THEORY

HOWARD LEVENE¹

Departments of Mathematical Statistics and Zoology, Columbia University, New York City

Received February 17, 1958

THIS is the fourth of a series of articles on the release of genetic variability through recombination. Identical experiments were performed on three species of *Drosophila*, and the results were given and analyzed in papers on *D. pseudoobscura* (SPASSKY *et al.* 1958), *D. persimilis* (SPIESS 1959) and *D. prosaltans* (DOBZHANSKY *et al.* 1959). The present paper describes the statistical methods and concepts used in all three analyses, discusses the definition and interpretation of the components of variance involved, and discusses some conceptual difficulties connected with the use of statistical tests of significance on material of this kind.

Model for the analysis of variance

The mathematical model for the analysis of variance of viability for the inter-population cross is as follows. Let X_{ijk} be the observed viability of the k th derived chromosome from the cross between the i th chromosome of population T and the j th chromosome of population C, where the populations are labelled T and C for definiteness. We suppose that X_{ijk} is the sum of a number of components:

$X_{ijk} = m + t_i + c_j + d_{ij} + w_{ijk} + r_{ijk} + b_{ijk}$, where m is the grand mean and all the other contributions are random effects with zero means and with variances $\sigma_t^2, \sigma_c^2, \sigma_d^2, \sigma_w^2, \sigma_r^2$ and σ_b^2 respectively.

Since the means are zero, if any of these variances is equal to zero the corresponding contribution will be identically equal to zero, and will play no part in determining the value of X_{ijk} . In explaining the meaning of these components it will be convenient to suppose that all components to the right of the one being discussed are equal to zero.

The component t_i represents the additive effect of the i th chromosome from population T, and has the same value for all chromosomes derived from the i th chromosome from population T. Similarly c_j represents the additive effect of the j th chromosome from population C, and has the same value for all chromosomes derived from the j th chromosome from population C. Thus at this level, the value of X_{ijk} can be completely predicted if we know the grand mean and the additive effects of the chromosomes from the two populations. The component d_{ij} is the nonadditive effect of the cross between the i th chromosome from population T,

¹ This research was sponsored by the Office of Naval Research under Contract Number Nonr-266 (33), Project Number 042-034.

and the j th chromosome from population C, and has a specific value for all chromosomes derived from this cross. Thus, as soon as this component is present, it is no longer possible to estimate the viability of a derived chromosome from the additive effects of the two chromosomes from which it was derived; and in fact d_{ij} measures the departure of the actual viability from this predicted value. The component w_{ijk} is the genetic component within crosses, and represents the difference between the actual viability of the k th derived chromosome from this cross and the mean viability characteristic of chromosomes derived from that cross. Thus σ_w^2 represents the variability released by recombination and is the primary component of interest in the present papers. The sum of the components through w then gives the true genetic viability of a particular derived chromosome, and in principle can be estimated to any desired accuracy by repeated measurements on the same derived chromosome. The component r_{ijk} is the component for the real residual variance and represents the difference between the true genetic viability of a given derived chromosome and its actual viability as affected by the residual genetic variance and, more important, environmental variance. Finally the component b_{ijk} represents the binomial sampling variance due to the fact that we only observed 100 flies.

The variances of the components just discussed are estimated in the components of variance column of Table 3. The components through d_{ij} can be estimated directly from the data on the interpopulational cross, but this data will only give us the sum of the last three components. However, the binomial component is a simple function of the grand mean, and thus can easily be estimated, and the component r_{ijk} includes just those factors affecting the replicate cultures of the original chromosomes, and thus can be estimated from that data.

For certain statistical procedures used it is necessary to assume that the variables entering into X_{ijk} are normally distributed. Now b_{ijk} is a binomial variable and is essentially normally distributed, and r_{ijk} , which mainly represents the effect of environment, is probably likewise nearly normally distributed. Furthermore, it seems reasonable to assume that w_{ijk} is normally distributed, except that there are too many extremely small values of w , corresponding to lethality. Again t_i , c_j and d_{ij} can probably be taken to be normally distributed without too much error. Finally X_{ijk} , being the sum of all these components is likely to be still more nearly normally distributed, except that the value zero, denoting lethality, occurs more often than would be expected if X were normally distributed. Eliminating lethals, the X_{ijk} observed are normally distributed to a sufficiently good approximation. Thus the analysis of variance might be more valid if the data were analyzed with lethals eliminated; however, reasons have been given in a previous paper of this series (DOBZHANSKY *et al.* 1959) for preferring to deal with all variation simultaneously. Luckily the F test of the analysis of variance is insensitive to moderate departures from normality and it is felt that it may safely be used on the present data. Estimation of the components of variance themselves does not require any assumption of normality.

The mathematical models for intrapopulational crosses and for the original

chromosomes follow the same lines, and need not be given here in detail. However, it might be mentioned that if the additive effects of the chromosomes in the intrapopulational crosses are denoted as a_1 to a_{10} , each X_{ijk} includes *two* a 's, corresponding to the i th and j th chromosomes.

Analysis of variance of means

The analysis of variance for the original chromosomes is as follows. For a given locality let X_{ij} be the viability (number of wild type flies in 100 flies) of the j th replicate culture of the i th original chromosome, let M_i be the mean viability of the 6 cultures of the i th chromosome, and let $\bar{M}$ be the mean of the ten chromosomes. Table 1 gives the analysis of variance, which is an ordinary one-way

TABLE 1
Analysis of variance for original chromosomes

Effect	Sum of squares	d.f.	Mean squares	F ratio	Components of variance
Between chromosomes	$6[\sum M_i^2 - M^2/10]$	9	$\bar{A} = (\text{sum of sq})/\text{d.f.}$	$\bar{A}/\bar{U}$	$(\bar{A} - \bar{U})/6$
Real residual		50	$\bar{V} = \bar{U} - \bar{B}$		$\bar{V}$
Binomial		..	$\bar{B} = \bar{M}(100 - \bar{M})/100$		$\bar{B}$
Total residual	$\sum[\sum_j X_{ij}^2 - 10M_i^2]$	50	$\bar{U} = (\text{sum of sq})/\text{d.f.}$		$\bar{U}$

analysis of variance with extra lines for the "real residual" and binomial components. It should be noted that this and the other analyses of variance in the present paper are given in the customary form for such analyses. The sums of squares and mean squares in each table are greater by a constant factor (six for the present table) from the form judged simplest for the explanation in the first paper of the series. However the F ratios and components of variance are identical in both forms, and as the sums of squares and mean squares were not given in the tables in the previous papers, no confusion should result. To help construct the present tables, the letters $\bar{A}$, $\bar{C}$, etc. are used to denote mean squares which have no letter assigned to them in the earlier papers.

The analysis of variance for the intrapopulational crosses is as follows. Let X_{ijk} be the viability of the k th derived chromosome for the cross between chromosomes i and j , and M_{ij} be the mean viability for this cross. Let M_i be the sum of the M_{ij} for the nine crosses involving the i th chromosome, $M_{..} = \sum M_i$, which is twice the sum of the 45 M_{ij} , let $\bar{M}_i = M_i/9$ be the mean viability for the crosses involving the i th chromosome, and let $\bar{M} = M_{..}/90$ be the grand mean. Also let L^2 equal the sum of the squared values of the 45 M_{ij} ; that is $L^2 = \sum \sum_{i < j} M_{ij}^2$. Also let $U_{ij} = \sum (X_{ijk} - M_{ij})^2/9$ and let $\bar{U}$ be the mean of all 45 U_{ij} . Finally, quantities like $\bar{U}$ derived from the cross in question will be denoted by the sub-

script $\bar{U}_{intra}$, while corresponding quantities calculated from the data for the *original chromosomes* will be denoted by $\bar{U}_{orig}$. The analysis of variances is given in Table 2. The factor by which the sums of squares and mean squares in this table exceed those discussed in the previous papers is 80. This analysis of variance, except for the lines for the real residual and binomial components, is of the type called a diallel analysis of variance, first discussed by SPRAGUE and TATUM (1942) and corrected by COCHRAN (FEDERER 1951). It is conveniently summarized by GRIFFING (1956).

The analysis of variance for the interpopulational crosses is as follows. For definiteness denote the two populations T and C. Then let X_{ijk} be the viability of the k th derived chromosome for the cross between the i th chromosome from T and the j th chromosome from C, and let M_{ij} be the mean viability of the ten derived chromosomes from this cross. Further let $M_{i.} = \sum_j M_{ij}$ and $M_{.j} = \sum_i M_{ij}$. Then $\bar{M}_{i.} = M_{i.}/10$ is the mean viability for all crosses involving the i th chromosome from T and $\bar{M}_{.j} = M_{.j}/10$ is the mean viability for all crosses involving the j th chromosome from C. Also let $M_{..} = \sum_i \sum_j M_{ij} = \sum_i M_{i.} = \sum_j M_{.j}$ be the sum of all 100 M_{ij} , and $L^2 = \sum_i \sum_j M_{ij}^2$ be the sum of all 100 M_{ij}^2 . Then the analysis of variance is given in Table 3. In this table quantities like $\bar{V}_{orig}$ are the *means* of the corresponding values for the original chromosomes in both populations. The factor by which the sums of squares and mean squares in this table exceed those discussed in the previous papers is 100.

The analysis of variance of variances

The model for the analysis of variance of means assumes that the component w_{ijk} has a constant variance σ_w^2 regardless of the cross in question, and that the estimates which could have been obtained individually for each cross would differ only because of sampling error due to the fact that only ten derived chromosomes are studied for each cross. In point of fact it seems more reasonable that different crosses involving different genetic material should also differ in the amount of genetic variability they release. Such variability in the value of σ_w^2 , if not too great, does not seriously interfere with the actual carrying out of the analysis of variance for means; nevertheless, it is of some interest to consider this variability for its own sake, and that is done in the present section. The sums of squares for these analyses of variance are obtained exactly as before, except that the values V_{ij} are used instead of the M_{ij} and the factor six in Table 1 and ten in the other two tables, by which the sum of squares is multiplied, is omitted. Similarly, the denominator in the corresponding component of variance is reduced by a factor of six in the first table and ten in the other two tables. The reason for this is that each V_{ij} is a single value whereas M_{ij} is already the mean of either six or ten observed values. Because each cross gives only a single value of V_{ij} it is not possible to obtain an empirical estimate of the variability of the variance within a cross. Nevertheless, it is essential to have some estimate of this variability if we are to tell whether there is a real nonadditive component between crosses or

TABLE 2
Analysis of variance for viability of derived chromosomes for intrapopulational crosses

Effect	Sum of squares	d.f.	Mean squares	F ratios	Components of variance
Additive between	$10[(\sum M_i^2)/8 - M^2/80]$	9	$\bar{C} = (\text{sum of sq})/\text{d.f.}$	$\bar{C}/\bar{D}$	$(\bar{C} - \bar{D})/80$
Nonadditive between	$10[L^2/2 - (\sum M_i^2)/8 + M^2/144]$	35	$\bar{D} = (\text{sum of sq})/\text{d.f.}$	$\bar{D}/\bar{U}_{intra}$	$(\bar{D} - \bar{U}_{intra})/10$
Within		405	$\bar{U}_{intra}$	$\bar{U}_{intra}/\bar{U}_{orig}$	$\bar{U}_{intra} - \bar{E}$
Real residual		50	$\bar{V}_{orig}$		$\bar{V}_{orig}$
Binomial		..	$\bar{B}_{intra} = \bar{M}(100 - \bar{M})/100$		$\bar{B}_{intra}$
Total residual		50	$\bar{E} = \bar{V}_{orig} + \bar{B}_{intra}$		$\bar{E}$

TABLE 3
Analysis of variance for viability of derived chromosomes for crosses between two populations designated T and C

Effect	Sum of squares	d.f.	Mean squares	F ratios	Components of variance
Additive between for T	$10[(\sum M_{i.}^2)/10 - M_{..}^2/100]$	9	$\bar{G} = (\text{sum of sq})/\text{d.f.}$	$\bar{G}/\bar{S}$	$(\bar{G} - \bar{S})/100$
Additive between for C	$10[(\sum M_{j.}^2)/10 - M_{..}^2/100]$	9	$\bar{H} = (\text{sum of sq})/\text{d.f.}$	$\bar{H}/\bar{S}$	$(\bar{H} - \bar{S})/100$
Nonadditive between	$10[L^2 - (\sum M_{i.}^2)/10 - (\sum M_{j.}^2)/10 + M_{..}^2/100]$	81	$\bar{S} = (\text{sum of sq})/\text{d.f.}$	$\bar{S}/\bar{U}_{inter}$	$(\bar{S} - \bar{U}_{inter})/10$
Within		900	$\bar{U}_{inter}$	$\bar{U}_{inter}/\bar{U}_{orig}$	$\bar{U}_{inter} - \bar{W}$
Real residual		100	$\bar{V}_{orig}$		$\bar{V}_{orig}$
Binomial		..	$\bar{B}_{inter} = \bar{M}(100 - \bar{M})/100$		$\bar{B}_{inter}$
Total residual		100	$\bar{W} = \bar{V}_{orig} + \bar{B}_{inter}$		$\bar{W}$

whether the observed mean square between crosses is just what might be expected due to sampling variability. If we can assume that the underlying data from which the V_{ij} are calculated (the X_{ijk}) are normally distributed we can obtain a theoretical value for this variance as follows.

If $X_1, \dots, X_N$ are independently normally distributed with mean m and variance σ^2 , then the sample variance s^2 has $N-1$ degrees of freedom and has mean σ^2 and approximate variance $2(\sigma^2)^2/(N-1)$, or, still more approximately, $2(s^2)^2/(N-1)$. In the present case, then, the theoretical variance of U_i for an original chromosome is $2U_i^2/(6-1)=2U_i^2/5$, while the variance of U_{ij} for a cross is $2U_{ij}^2/(10-1)$. Now $V_{ij}=U_{ij}-B_{ij}$ has variance approximately $2U_{ij}^2/9 + \text{Var}(B_{ij})$, since B_{ij} depends only on M_{ij} and can be considered statistically independent of V_{ij} . Further, since B_{ij} is estimated from 1000 flies, $\text{Var}(B_{ij})$ is negligible compared to $2U_{ij}^2/9$, and we have the variance of V_{ij} estimated by $2U_{ij}^2/9$. Finally, a suitable estimate of the variance of the V_{ij} based solely on variability within crosses is $2\bar{U}^2/9$, where $\bar{U}$ is the mean for the set of crosses in question. This estimate is entered as the "within" mean square and used in calculating the nonadditive between crosses F ratio and component of variance. Its use for the F ratio is not strictly correct, since $2\bar{U}^2/9$ is not distributed as a mean square, but it is based on so many more degrees of freedom than the numerator that the distribution of F is not seriously affected. Similar considerations apply to the original chromosomes, where $2\bar{U}^2/5$ is entered as the total residual mean square. For the crosses, a total residual mean square has been entered as $2\bar{U}_{orig}^2/9$, and used to obtain an estimated component of variance within crosses. These last entries serve no useful purpose except convenience in comparison; they give no information not already contained in the corresponding components for the means.

Nonparametric tests

It has been pointed out above that the F test requires the assumption of normality, although it is not very sensitive to moderate departures from this assumption. Even if the basic data, X_{ijk} , are normally distributed, the variances calculated from them, the U_{ij} , will not be normally distributed. However, the more degrees of freedom there are for an estimated sample variance the more nearly it is normally distributed. In the present case for the crosses the U 's have nine degrees of freedom, and there is some evidence that this is sufficient to allow for the use of the analysis of variance. Since the binomial component has only a small variance, the V_{ij} are essentially distributed as the U_{ij} and we may use the analysis of variance on the V 's without too much worry. Nevertheless, it might be preferable to use a test that did not require the assumption of normality. In the case of the V_{ij} such a test is available only for testing the additive effect. For the inter-populational crosses the test of FRIEDMAN (1937) also described in KENDALL (1948) may be used, while a modification of this test can be used for the diallel analysis of variance. These tests are based on ranks, and somewhat greater sensitivity can be obtained if these ranks are transformed into the equivalent normal deviate by the use of Table 20 in FISHER and YATES (1938). The F values for the

additive effect for the variances in Table 5 of each of the first three papers of this series were obtained by these nonparametric methods.

The interactions of interactions

The analysis of variance of variances was undertaken to investigate the question of whether σ_w^2 was constant or whether different w_{ijk} 's had different variances depending on the particular cross involved. Similarly, one can ask whether σ_d^2 is constant or whether the variance of d_{ij} depends on i and j (but not on the specific combination of the two); i.e., $Var(d_{ij}) = l_i + q_j$, say. In other words we may ask whether certain chromosomes have a tendency to react erratically, with a large nonadditive component, whereas other chromosomes are more consistent and the viability of chromosomes derived from crosses in which they take part may be predictable. The method used for testing this will be described more fully elsewhere. Here it will suffice to say that the individual d_{ij} were estimated by the difference between the mean viability for a particular cross and that predicted from the grand mean and the additive effects of the chromosomes involved, and that a nonparametric analysis of variance was done on the absolute values of these estimated interactions. The results are given in Table 5 of each of the first three papers in the lines labelled int. of int., for interaction of interactions.

Confidence limits for the components of variance

Confidence limits for a single mean square may easily be obtained by the use of a table of chi-square, and this method is explained, for example, in DIXON and MASSEY (1957). The genetic components of variance for the means and the components of variance between crosses for the variances are estimated by a difference of two mean squares. Confidence limits for these have been obtained by the use of the second approximation of BULMER (1957), with slight modification for the case of the component of variance within crosses for the means. For the means, confidence limits for the total residual component are obtained from the confidence limits for $\bar{U}$ for the original chromosomes with a slight modification. Confidence limits for the binomial coefficient are obtained from confidence limits for the grand mean for the cross. Finally the confidence limits for the real residual variance are obtained by subtracting the binomial estimated component from the limits for the total residual component. It should be noted that the lower limit for the total residual component is never allowed to be less than the lower limit for the binomial component, since the total residual variance is the sum of the binomial and real residual components.

Unlike the F test for the existence of the various components of variance, the method used for obtaining confidence limits for these components is extremely sensitive to departures from normality in the data; consequently the limits as given can only be considered as very rough approximations giving the order of magnitude of our uncertainty in knowledge of the components. Whenever the F test in Table 5 of any of the papers showed that a certain component of variance was not significantly different from zero at the $21\frac{1}{2}$ percent level of significance,

the corresponding lower 95 percent confidence limit for that component was set equal to zero in Table 6. On the other hand in one case for *Drosophila persimilis* the lower confidence limit of zero was obtained although a nonparametric F test showed the existence of a significant component. Since the nonparametric F test cannot be used to estimate components of variance, all that could be done was to indicate in the text that there was probably a significant component but that it could not be estimated.

The meaning of significance tests

When we say that a certain effect is significantly different from zero at the five percent level, we mean that if there were no true effect then we would obtain as large an estimated effect in less than five percent of all repetitions of the experiment. Similarly a standard error attached to an estimate is a measure of the variability of that estimate in repeated experiments on the same material. However, in many cases, the meaning of the words "repeated experiments" are by no means clear. For example, in the case of the present material, we can conceive of at least five kinds of repetitions of the experiment. (1) We can use the same chromosomes and crosses, repeated at the same time the original experiment was done. (2) We can repeat the experiment using the same chromosomes and crosses but at a somewhat later time, when the genetic contents of the chromosomes may have changed somewhat because of mutation, and experimental conditions will certainly be somewhat different. (3) We can repeat the experiment with other chromosomes taken at random from the same population. (4) We can repeat the experiment using other populations of the same species. (5) We can repeat the experiment using different species.

Each of these types of repetition involves all the sources of variation present in the previous kinds of repetition plus some additional ones, and consequently the further along we are in the series the greater the amount of variability that would be observed.

Since for each original chromosome six simultaneous replicates were tested, and for each cross ten derived chromosomes were tested simultaneously, we have available a repetition of type (1). The most important problem in the present papers is the comparison of the variability released by recombination within crosses with the variability between replicate cultures for the same original chromosomes. Here repetition of type (1) is appropriate and it is this repetition that furnishes the error estimates used in the analyses of variance and in the confidence intervals for the variance components. Since different crosses were tested at various times over a period of a year or more, and the original chromosomes were tested at still another time, a comparison of the variability between crosses with that within crosses should properly involve repetition of type (2). No repetition of this type is available, but the additional variability should be rather small, and thus the use of type (1) repetition should be satisfactory. It is true that for one locality, Texas, in *Drosophila pseudoobscura*, the original chromosomes were tested at two different times. However, these times were even

further apart and the conditions of the tests were somewhat different, so that this would give us an overestimate of the amount of additional variability in repetition of type (2).

If it is desired to make comparisons between the two localities for a given species still further problems arise. If we are simply interested in whether there are real differences between the two particular groups of chromosomes studied, then the proper error is obtained from the mean square within crosses. However, if we are really interested in whether there are differences between the localities *per se*, then we are interested in repetition of type (3); that is, the question is what would happen if we obtained new samples from these two localities. In this case the error must be obtained from the mean square between crosses, which is considerably bigger, and furthermore is derived from fewer degrees of freedom and consequently is less likely to yield significant differences. A further complication here is that the chromosomes used are not a random sample from the population in question, but were selected to have moderately high viability when homozygous. Consequently the conclusions are only valid for the class of randomly selected chromosomes of high viability. Finally, in comparing the behavior of different species, a further problem of differences between localities arises. If we can assume that all localities are the same, which is very doubtful, then replication of type (3) is proper and comparisons between species are made in precisely the same way as comparisons between localities for the same species. If, however, we make the more realistic assumption that there are differences between different localities for the same species, then repetition of type (4) is in order; that is we must compare the results from a random sample of localities for one species with the results from a random selection of localities for the second species. As, in the present case, we have only two localities for each species, it is very difficult to prove any differences between species.

The above, while rigorously correct, takes, perhaps, a somewhat pessimistic view of the matter. It is not always necessary to have results that are statistically significant at the five percent level. In some cases it may be enough if they are significant at the 50 percent level; that is, there would be only an even chance of observing such a difference if in fact no true difference existed. Furthermore, one can always make estimates of various effects and various differences, and it is not necessary to assume that such estimates will always differ from their true values by, for example, two standard deviation units. In fact most of the time they will differ by much less. Thus with the present data it would seem reasonable, for example, to say that there was a difference in the amount of variability released between Pirassununga and Rio, for *Drosophila prosaltans*, and also that there was a difference between the amount of variability released in *Drosophila pseudoobscura* and *Drosophila persimilis*, even though in neither case can the difference be rigorously proved statistically. In this connection it should be realized that even if perfectly correct statistical methods yield a result that is statistically significant at the one percent level, it is still possible that no effect actually exists and that this is just one of the one percent of the cases in which an

abnormally large observed effect is found. Thus, the difference between effects that are proved statistically in a rigorous fashion and those that are only shown in a somewhat looser way is not necessarily a hard and fast distinction.

It is evident from the above that a given single estimate, such as, for example, a mean viability, might have four different standard errors depending on the use to which the standard error was to be put, and attaching a single standard error could prove quite misleading. Accordingly, no standard error has been attached to such estimates, with the exception of a few cases where for some reason it was felt desirable to do so.

One other question of significance that arises is that of correlations. All correlations in these papers are based on ten chromosomes. If we are interested only in the correlation between these particular chromosomes, most of the coefficients observed are significant. If, however, we wish to make inferences about what would happen with ten other chromosomes from the same population, then the observed correlation must be less than -0.63 or greater than $+0.63$ to be significant at the five percent level. Evidently ten observations are too few to give much information about a correlation. Nevertheless the correlations are included, since the presence of many small ones and a number of negative ones, shows that in general there is little consistency in the contribution of a given chromosome under different conditions.

SUMMARY

This paper discusses the statistical methods used in earlier papers of this series to analyze data on *Drosophila pseudoobscura* (SPASSKY *et al.* 1958), *D. persimilis* (SPIESS 1959), and *D. prosaltans* (DOBZHANSKY *et al.* 1959). The use of diallel analysis of variance and two-way analysis of variance, combined with a one-way analysis of variance on control data, to test hypotheses about various components of variance, and to estimate these components and obtain approximate confidence intervals for them is explained. The definitions of the various components in terms of the mathematical model used and their genetic interpretation are given. This is done for components of variance of variances as well as for the more usual components of variance of means. Certain difficulties inherent in the interpretation of significance tests, confidence limits, and standard errors for complex experiments such as the present are discussed.

LITERATURE CITED

- BULMER, M. G. 1957 Approximate confidence limits for components of variance. *Biometrika* 44: 159-167.
- DIXON, W. J., and F. J. MASSEY, JR., 1957 *Introduction to Statistical Analysis*. McGraw-Hill. New York.
- DOBZHANSKY, TH., H. LEVENE, B. SPASSKY, and N. SPASSKY, 1959 Release of genetic variability through recombination. III. *Drosophila prosaltans*. *Genetics* 44: 75-92.
- FEDERER, W. T., 1951 Evaluation of variance components from a group of experiments with multiple classifications. *Agr. Exp. Sta. Iowa State College. Research Bull.* 380.

- FISHER, R. A., and F. YATES, 1938 *Statistical Tables for Biological, Agricultural and Medical Research*. Several editions. Oliver and Boyd. London.
- FRIEDMAN, M., 1937 The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statistical Assn.* **32**: 675-701.
- GRIFFING, B., 1956 A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity* **10**: 31-50.
- KENDALL, M. G., 1948 *Rank correlation methods*. Griffin. London.
- SPASSKY, B., N. SPASSKY, H. LEVENE, and TH. DOBZHANSKY, 1958 Release of genetic variability through recombination. I. *Drosophila pseudoobscura*. *Genetics* **43**: 844-867.
- SPIESS, E. B., 1959 Release of genetic variability through recombination. II. *Drosophila persimilis*. *Genetics* **44**: 43-58.
- SPRAGUE, G. F., and L. A. TATUM, 1942 General vs. specific combining ability in single crosses of corn. *J. Am. Soc. Agron.* **34**: 923-932.