THE ANALYSIS OF TETRAD DATA [1, 2]

H. P. PAPAZIAN
Department of Botany, University of Chicago, Chicago, Illinois

Received June 21, 1951

THE theory of four-strand crossing over was founded by JANSSENS (1909). He arrived at his theory in an attempt to explain tetrads in which all four cells are different. None such had been observed at that time but JANSSENS believed, by a happy metaphysical insight, that since germ cells were always produced in groups of four, the four must be all different. It is curious that when tetrads were isolated and four-type tetrads obtained, their significance was often missed. A useful summary of the types of tetrad produced by various crossovers has been given by WHITEHOUSE (1942, 1949).

Haploid characters among the bryophytes, green algae, and fungi have been used almost exclusively in tetrad analyses although flowering plants would seem to offer excellent material. In many flowering plants, including nearly all the Ericaceae, pollen is shed in tetrads. If individual tetrads were placed on styles there should develop, with or without hormone treatment, four seeds corresponding to the four pollen grains of the tetrad.

The purpose of the present paper is to show what information can be obtained from tetrad analysis and to analyze some recent published data.

Tetrads can be classified into ordered and unordered kinds. In ordered tetrads the pairs of cells resulting from the first meiotic division can be distinguished and hence the second division segregation frequency for any gene pair estimated. In organisms which form spores in a linear order such as many ascomycetes and the smuts, they are represented by the top two or four spores and the bottom two or four. In organisms in which the tetrads form a square and in which the second meiotic division spindles form a cross, they are represented by spore pairs along the diagonals, which is probably the case in chiastobasidial basidiomycetes. Another type of ordered tetrad is represented by *Saccharomycoides ludwigii*. Here the second meiotic division spindles are parallel and at about the same level so that one daughter cell from each goes to the top and one daughter cell from each goes to the bottom (WINGE 1947). The top and bottom pairs of cells are not sister cells but are, so to speak, cousin cells. In this case the second division segregation frequency for a locus is twice the frequency of tetrads in which the top pair and bottom pair of cells have similar alleles.

In unordered tetrads spore pairs cannot be distinguished by their spacial position. In some cases they can be distinguished by other means. In organ-

[1] This work was supported by a grant from the Dr. Wallace A. and Clara A. Abbott Memorial Fund of the University of Chicago.

[2] The mathematical appendices were done with the assistance of L. J. Savage of the Committee on Statistics, University of Chicago.

isms having a sex chromosome sex provides a convenient marker where it is known that the sex locus is always pre-reduced, this method has been used in *Sphaerocarpus* (KNAPP 1937). A gene so close to the centromere that practically no crossing over takes place between it and the centromere would also provide such a marker. The extra information that can be obtained from ordered tetrads or from unordered tetrads with a marker concerns crossing over in intervals bounded by a centromere. Information about the position of the centromere can, however, be obtained from any kind of tetrad so long as at least three gene loci are available.

One of the advantages of tetrad analysis over the analysis of random spores is, then, that it gives information about the position of the centromere. Another advantage is that it provides more precise information about interference. With tetrads interference can be estimated with as few as two gene loci, and when more loci are available chromatid interference can be distinguished from chiasma interference. With random spores chromatid interference can only be detected by the exceptional occurrence of recombination values in excess of 50 percent. Finally, in multi-loci crosses the frequencies of crossovers in each interval and of simultaneous crossovers in two or more intervals can be directly estimated. Indirect estimates of these can also be made with random spores (LUDWIG 1938; WEINSTEIN 1936) if chromatid interference is assumed absent. For most purposes these estimates are fortunately not required but for some they are (CHARLES 1938).

When linkage is estimated from recombination frequencies the collection of spores in tetrads is less efficient than random collection. If R is the frequency of recombination per strand between two linked loci, the variance of the estimated linkage value V_r, is $[R_r(1-R_r)]/n_r$ where n_r is the total number of spores collected randomly. When the spores are collected in tetrads the variance of the estimated linkage value V_t is $(R_t - 2R_t^2 + N)/2n_t$ where n_t is the total number of tetrads and N is the frequency of tetrads having all four spores showing recombination. This formula is equivalent to that given by MATHER and BEALE (1942) and L. J. SAVAGE has pointed out to me that although these authors arrived at this formula by an approximate calculation, an exact calculation gives the same result. If data from random spores and from tetrads are to be pooled the results from each should be weighted by the reciprocal of their variances so that the pooled estimate of recombination is $[R_r(1/V_r) + R_t(1/V_t)]/[1/V_r + 1/V_t]$ or $(R_rV_t + R_tV_r)/(V_r + V_t)$. Two random spores give about as much information as all four spores of a tetrad since the variance of R_t is about twice that of R_r. An example of the use of such weighting is provided by data from *Neurospora* (SHENG 1951) where linkage between the loci ***nd*** and ***al***$_2$ was estimated from 172 random spores and 24 tetrads. The four spores of a tetrad were apparently given the same weight as one single random spore to arrive at the mean value 20.5. By the above method $V_r = 0.000961$, $V_t = 0.00385$ and the mean value is 19.6. This disadvantage of tetrads is mitigated by the fact that if three spores of a tetrad are known the fourth can always be deduced and if only two spores

are known but those two have similar alleles for all but one of the characters involved, then the other two can be deduced.

CLASSIFICATION OF TETRADS

In the following discussion the centromere will be considered as a locus comparable to a gene locus. From this point of view unordered and ordered tetrads can be analyzed by similar methods. Pairs of loci will be classified into two-type parental (P), two-type non-parental (N), and four-type (F) in which there are 0, 2, and 4, recombinants in the tetrad respectively. This classification reduces the number of classes and simplifies analysis. Where one of the loci of a pair is the centromere then P and N cannot be distinguished and F represents the incidence of second division segregation of the gene locus. Whereas with random spores, there is only one independently variable frequency for each pair of loci, parentals or non-parentals, with tetrads there are two independently variable frequencies. It is often useful to know the number of classes of tetrad that can be expected from a cross with a certain number of factors, say x. The total number of possible arrangements of the four spores resulting from meiosis is 6^x. The number of classes of ordered tetrads, pooling only those classes that give no information as to crossing over is $(5 \times 2^x + 6^x)/8$. The number of classes of unordered tetrads $(6^{x-1} + 3 \times 2^{x-1})/4$, and if all pairs of loci are classified as N, P, or F, then the total number of classes is a bit more difficult to calculate but for the first ten values of x the series is 1; 3; 11; 48; 236; 1,248; 6,896; 39,168; 226,496; 1,325,577 (appendix I). The classification into N's, P's, and F's reduces the number of classes to be distinguished considerably. The information lost is, at worst, a failure to distinguish between the two types of 3-strand double crossovers (PAPAZIAN 1951).

With this scheme the detection of crossovers and the strands involved is simplified. For any three loci the types of crossovers corresponding to various combinations of N's, P's and F's are given in table 1. In the first three columns if an N occurs instead of a P, a four-strand double crossover is indicated instead of no crossover. It is impossible to have an N for all three pairs of loci so that by substituting N's for P's only three additional classes can be added to the first column and one additional class to the second and third column.

TABLE 1

Classes of tetrad produced by various kinds of crossing over.

Crossovers	None	Single in I	Single in II	Double, in I and II		
				2-str.	3-str.	4-str.
Loci pairs: A & B	P	F	P	F	F	F
B & C	P	P	F	F	F	F
A & C	P	F	F	P	F	N

The three loci lie in the order A–B–C with two intervals I and II respectively. P = 2-type parental, N = 2-type non-parental, F = four-type.

The table then accounts for the 11 possible classes with three factors. Where one locus is the centromere, for example the *A* locus, and the two gene loci lie in the same arm of the chromosome, two- and four-strand double crossovers cannot be distinguished since P and N cannot be distinguished where a centromere locus is involved. Similarly, if the center locus, *B*, is the centromere and a four-strand double crossover occurs, it is not possible to tell in which interval it occurred. In this case the A–C pair will be N and one of the other two pairs will be N or P but which is which will be unknown. Any number of loci can be analyzed in this way in groups of three but a complication arises when there are two three-strand double crossovers. With four loci, for instance, if there are three-strand doubles in intervals I & II and in II & III it is useful to know whether the last crossover involved the same strands as the first. In the former case the first and last locus pair will be P, in the latter case N.

FREQUENCIES OF THE CLASSES F, N, AND P

If the number of crossovers occurring simultaneously in an interval is n, and it is assumed that there is no chromatid interference, the proportion of F's between the bounding loci is given by $F = (2/3)[1-(-1/2)^n]$. This formula was derived by MATHER (1935) for the frequency of second division segregation but in the scheme proposed here it is valid for F's between two gene loci as well as between a gene- and a centromere-locus. To proceed further some assumption must be made about chiasma interference. It is convenient to consider two models as reference points, that of no interference in which multiple crossovers occur in a Poisson distribution, and that of complete interference in which only single crossovers occur. If the mean number of crossovers in an interval is m, then with complete interference the frequency of recombinants, $R = m/2$, and $F = m$; hence $R = F/2$. Also $N = 0$. With no interference it is well known that

$$R = (1/2)(1-e^{-m}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

and

$$F = (2/3)(1-e^{-3m/2}) \text{ (PAPAZIAN 1951)} \quad \ldots\ldots\ldots \quad (2)$$

but

$$R = N + F/2$$

so from (1) and (2)

$$N = 1/2 - (1/3)(1-e^{-3m/2}) - e^{-m}/2 \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

and since $F + N + P = 1$

$$P = 1/2 - (1/3)(1-e^{-3m/2}) + e^{-m}/2 \quad \ldots\ldots\ldots\ldots\ldots \quad (4)$$

From (2) and (3)

$$2N = 1 - F - (1-3F/2)^{2/3}$$

$$N = F^2/8\,(1+2F/3) \text{ approximately} \quad \ldots\ldots\ldots\ldots\ldots \quad (5)$$

and

$$R = F/2 + F^2/8\,(1+2F/3) \text{ approximately} \quad \ldots\ldots\ldots \quad (6)$$

Departures from a Poisson distribution of exchanges or from random strand multiple exchanges can be detected by departures from (5) and (6) where only two loci are available. This is valid for long intervals and can be used where other methods of measuring interference cannot. As map distances increase $m \to \infty$ and $F \to 2/3$, $N \to 1/6$, $P \to 1/6$. These limiting values obtain for two independent or loosely linked gene loci one or both of which are far from the centromere. Values of F in excess of 2/3 indicate either an excess of single crossovers or an excess of 3-strand doubles. Values of R in excess of 1/2 always indicate an excess of 4-strand doubles.

THE ADDITION LAW AND THE POSITION OF THE CENTROMERE

Considering recombination frequencies, if the mean number of crossovers, m, is taken to be additive, *i.e.*, $m_{1,2} = m_1 + m_2$, then from (1) (preceding section) $R_{1,2} = R_1 + R_2 - CR_1R_2$ where C varies from 0 with complete interference to 2 with no interference (HALDANE 1919). For F frequencies the corresponding addition law is, from (2) $F_{1,2} = F_1 + F_2 - C_F F_1 F_2$ where C_F varies from 0 to 3/2. In the case of two gene loci on different chromosomes absence of interference may be assumed and since all centromeres undergo first division segregation they can all be regarded, for the present purpose, as a single locus the relationship $F_{1,2} = F_1 + F_2 - (3/2)F_1F_2$ holds. The two chromosomes can be thought of as joined at their centromeres and $F_{1,2}$ the total distance between the two gene loci is the sum of each F distance between gene and centromere. This formula is identical with that given by PERKINS (1949) and WHITEHOUSE (1949) following a different line of argument.

When there are as many as three unlinked loci it is possible, by means of this formula, to map each locus in relation to its centromere (PERKINS 1949; LINDEGREN 1949; WHITEHOUSE 1950). The $F_{1,2}$ values of each of the three pairs of loci are equated to the corresponding F_1 and F_2 values which are unknown and since there are three equations and only three unknowns they can be solved. The solutions are of the form

$$F_1 = 2/3 \pm (2/3)\sqrt{(2-3F_{1,2})(2-3F_{1,3})/2(2-3F_{2,3})}.$$

It is clear that where F_1 has two solutions, their sum will always be 4/3. Not all combinations of alternative solutions will, however, be valid when they are tested in the original equations. It can, in fact, be shown that it is impossible for more than two sets of solutions ever to be valid. The solutions are unreliable for F_1 when $F_{2,3}$ has a value near 2/3.

Even where there are two loci on one chromosome and one on another chromosome the position of the centromeres can be estimated (PAPAZIAN 1951). In this case there will be two equations of the type $F_{1,2} = F_1 + F_2 - (3/2)F_1F_2$ for the pairs of loci on different chromosomes but for the pair of loci on the same chromosome absence of interference cannot be assumed and the constant C_F cannot be assumed to be 2/3; for short distances zero is a better approximation. Where the two linked loci are on different sides of the centromere the third equation is the same as before but where they are on the

same side a subtraction is required and the appropriate equation is $F_{1,2} = F_1 - F_2 + C_F F_1 F_2$.

EVIDENCE FOR INTERFERENCE FROM TETRAD DATA

Good evidence for chromatid interference and for negative chiasma interference in the sex chromosome of *Neurospora crassa* is given by WHITEHOUSE (1942) and LINDEGREN and LINDEGREN (1942) where references to cytological and genetical data bearing on these phenomena in other organisms can be found. In both of these references the same data from LINDEGREN were used in part.

There are a few cases of peculiar tetrad frequencies in the literature which, although open to various objections, deserve attention. WETTSTEIN (1924) describes a cross in the moss *Funaria hygrometrica* with four linked factors. Among 35 tetrads isolated all were two-type although crossing over (2-type non-parentals) was observed in all three intervals. Two-type tetrads were also obtained exclusively under certain conditions by MOEWUS in *Chlamydomonas* (1940), *Brachiomonas* (1940), and *Protosiphon* (1949). In the case of *Chlamydomonas* this was only true at low temperatures; at 22°C 4-type tetrads were obtained. Such results indicate either the regular occurrence of 4-strand double crossovers close together, or of crossing over at the two-strand stage.

In yeast (LINDEGREN 1949), among 83 unordered tetrads the proportion of N's between the loci PN and AD, situated on either side of the centromere, is 0.217. The value of F for the same two loci is 0.53 giving a calculated value of N of 0.042, so that the proportion of N's found is in great excess on an assumption of randomness. It is of interest to examine the apparent anomaly that the total distance between PN and AD is greater than the sum of their separate distances from the centromere. The former is estimated from recombination values and the latter are estimated from segregation values. Following the relations developed in the last section, assuming no interference, the expected value of R can be calculated. The distances of PN and AD from the centromere are 26 and 11 respectively. Therefore $F_1 = 0.52$, $F_2 = 0.22$ and, by the addition law for F's, $F_{1,2} = 0.5695$. From the relation $R = 1/2 + (F^2/8)(1 + 2F/3)$ the theoretical recombination frequency for an F frequency of 0.5695 is 0.33, or 33 map units. The actual map distance between PN and AD was found to be 48. The excess indicates an excess of double crossovers or of 4-strand doubles. The significance of these indications is only tentative owing to the fact that the linkage values of the five gene loci in these crosses are large, the number of asci analyzed relatively small, and the position of the loci therefore uncertain.

The numerous data published by HOULAHAN, BEADLE, and CALHOUN (1949) provide independent evidence on interference in the sex chromosome of *Neurospora crassa*. HOULAHAN *et al.*'s arrangement in table 6 was stated to be the most probable location, but to be based on inadequate data in some cases. Perkins has pointed out to me that crosses 7, 16, 17, and 36 (counting

from the top of table 1) involve genes known with considerable certainty to be in the arm opposite *sex*, whereas the genes in crosses 28, 47, 48, 52, 57, and 66 are less reliably located. Among the four crosses involving the more reliably located genes represented by 139 asci, the number of 2-, 3-, and 4-strand doubles are 2, 1, and 3 respectively. Although there is an excess of 2- and 4- over 3-strand, the numbers are too small to be significant. If the less reliably established loci are added, among 245 asci the numbers are 10, 2, and 4 in which case the excess of 2- and 4- over 3-strand is significant at the one percent level. This compares with a ratio of 20 : 10 : 5 (WHITEHOUSE 1942) and 32 : 14 : 13 (LINDEGREN and LINDEGREN 1942). In table 2 (HOULAHAN *et al.*) there are 14 crosses (Numbers 1–5, 7–15) involving sex and the albino 4637 which is linked to sex. There are 23 asci (= 0.0827) which are N for these two loci and 51 (= 0.183) that are F. The calculated proportion of F's is 0.00647. This probably indicates both an excess of 4-strand doubles and negative chiasma interference since even with the most favorable, and improbable, chiasma distribution where only double crossovers occur, the N's cannot exceed 1/4 the F's unless there is chromatid interference as well. It is noteworthy that in these crosses there are 22 asci showing double crossovers both in one interval and only one ascus with a double crossover, one in each interval. I am again indebted to D. D. PERKINS for pointing out that 4637 involves a translocation (MCCLINTOCK 1945). This would seem to be connected with the peculiar crossing over pattern but not in any way that is obvious. From similar crosses in table 1 (HOULAHAN *et al.*) the proportions of single crossovers in one arm, single crossovers in the other arm, and double crossovers, one in each arm was analyzed by a new statistical procedure devised by L. J. SAVAGE (appendix II). If genes which are reliably located (numbers 7, 8, 14, 16, 17, 22, and 36) are used, among 197 asci there is no significant excess or deficiency of doubles. If unconfirmed loci (numbers 28, 47, 52, 57, and 66) are added there is an excess of doubles significant at the one percent level. Double crossovers, both in the same arm, are ignored in these calculations. A possible explanation of negative chiasma interference is that crossing over is favored or inhibited in some asci of a perithecium by slight differences in environment or by accidental causes. If this were valid the same negative interference should be found for crossing over in different chromosomes. This was tested from data in table 3 (HOULAHAN *et al.*). Among 912 asci involving loci in different chromosomes there was no significant departure from independence. Again from table 1 (HOULAHAN *et al.*) among 883 asci there was no significant departure from independence of crossing-over between *sex* and the centromere and various other loci and their centromeres on other chromosomes. What departure there was was in the direction expected from a deficiency of simultaneous crossovers or positive interchromosomal interference.

Independent data from the loci sex and 35809 in *N. crassa* (MCDEVITT and BARATT unpublished; BUSS 1944) which lie 5.9 and 9.4 units on either side of the centromere indicate negative chromosome interference across the cen-

tromere. Among 85 asci, pooled from the two sources, there are 10 crossovers in the sex-centromere interval, 16 in the 35809-centromere interval, and four double crossovers, one in each interval. This is about twice the expected number of doubles on a basis of independence of the two events and by exact calculation of a 2×2 table is significant at the one percent level.

Two recent publications give some information as to how an increase or a decrease in total crossing over affects chiasma and chromatid interference. In *N. crassa* data are available (LINDEGREN and LINDEGREN 1942) for crosses made, accidentally, at two different temperatures. The mean frequency of crossovers between the extreme loci sex and pale was about the same at the high (h), and low (l) temperature being 0.404 and 0.390 respectively. At the high temperature the frequency of multiple crossovers was higher, 0.0688 (h) and 0.037 (l), which is significant at the one percent level. Owing to this the frequency of recombinants was actually higher at the low temperature 0.153 (h) and 0.168 (l). This effect was only evident in regions spanning the centromere. Within chromosome arms multiple crossovers were more frequent, but not significantly, at the low temperature 0.012 (l), 0.0089 (h). As regards chromatid interference, at the high temperature the proportion of 3-strand double crossovers was greater relative to the 2-strand and to the 4-strand doubles than at the low temperature, but the 2- and 4-strand doubles were still in excess over random expectation.

In *N. sitophila* it has recently been shown (FINCHAM 1951) that there is more crossing over in the region of the centromere than in *N. crassa*. The data in this paper are interesting and it is to be regretted that the classes of asci have been pooled to such an extent that detailed analysis is impossible. The proportion of F's between the centromere and *crisp* in *N. sitophila* is 0.711 which is in excess of 2/3 and suggests positive chiasma interference or an excess of 3-strand doubles. The proportion of F's between *crisp* and *sex* is 0.666 and the proportion of N's is 0.137 which is not in excess and does not contradict the possibility of an excess of 3-strand doubles or of positive chiasma interference. When F is nearly 2/3 the formula $2N = 1 - F - (3F/2)^{2/3}$ is very unreliable since a small change in F corresponds to a large change in N, the approximate formula given in an earlier section is still more unreliable when F is over 1/2 or so. It would therefore be difficult to tell whether there were a deficiency of N's between crisp and sex.

Independent data (WÜLKER 1935) from *N. sitophila* also indicate that in this species there is no excess of 4-strand doubles, and no negative chiasma interference as found in *N. crassa*, on the contrary, an excess of 3-strand doubles is suggested. The phenotypes of 180 asci analyzed for the factors *sex* and *l/L* (no aerial mycelium) are given in table 4 from which the following were calculated. The frequency of crossovers in the sex arm was 0.422, in the *l/L* arm 0.6, and in both arms simultaneously 0.24 which latter is just under the random value of 0.25. The frequency of F between sex and *l/L* was 0.678 and of N, 0.0833 which is under the expected random value of N

of 1/6. The number of 2-, 3-, and 4-strand doubles is 8, 28, and 8 respectively, this excess of 3-strand doubles is, however, only significant at about the seven percent level.

It would appear from the foregoing as if suppression of crossing over in *N. crassa* relative to *N. sitophila* were accompanied by a considerable relative increase in 2- and 4-strand double crossovers and also in a relative increase of multiple crossovers or negative chiasma interference.

THE SIGNIFICANCE OF CHROMATID INTERFERENCE

The effect of chromatid interference on the usual estimations of map distance and of chiasma interference cannot be treated quantitatively until a measure of chromatid interference is adopted. It is difficult, with present knowledge, to know what sort of measure would be suitable. A simple measure would be one analogous to chiasma interference and would assume that the chances of a particular strand being involved in a second crossover was less or greater than random. This would mean that with positive chromatid interference the departure from randomness would be in the direction 2-< 3-<4-strand and for negative interference 2->3->4-strand. Some of the data, on the other hand, suggest something different, namely that both 2- and 4-strand doubles may be in excess or in deficiency over 3-strand doubles.

The effects of any particular pattern of multiple crossovers can, however, be worked out. One interesting case is that of regular pairs of 4-strand crossovers. If, in such a case, chiasma interference were measured from random spores in the usual way, doubles in adjacent regions would not be detected at all. If the regions under consideration were not adjacent, and one crossover occurred in between, simultaneously with a double, then it would be detected always. Thus chiasma interference would apparently diminish with the distance of the regions concerned. With the usual assumption of chromatid randomness 1/2 of the doubles are detected since the frequency of single crossover strands is 1/2 the frequency of single exchanges and the frequency of double crossover strands is 1/4 that of double exchanges in the tetrad, irrespective of whether the regions are adjacent or not.

The cytological basis for chromatid interference is obscure. Just as chiasma interference can be accounted for, in a vague way, by most theories of crossing over, so chromatid interference can be fitted into many theories. On DARLINGTON's theory it might be due, as suggested by WHITEHOUSE (1942), to two particular strands being more liable to break than the other two. This would not account for an excess or deficiency of both 2- and 4-strand doubles over 3-strand doubles. Among other prominent theories of crossing over that of MATSUURA and HAGA (1942) not only might produce an excess of 2- and 4-strand doubles but would be expected to give a large excess. Furthermore, since the centromeres are post reduced in 2/3 of the cases all genes should show 2/3 post reduction unless a specific pattern of exchange between chromatids from different pairs occurred. On the theory proposed by BELLING

(1933) an excess of 2-strand might be expected. This theory differs from DARLINGTON's in its emphasis on the differentiation of the chromatid into genetic and non-genetic material, each of which duplicate independently. Single crossovers and multiple crossovers involving the same two strands are consequences of the hypothesis that non-genetic material " duplicates " by the formation of new strands independently, as regards position, of the old. Three- and 4-strand doubles involve, in addition, an exchange of strands of a different kind which might occur within the gene, or at the junction of genetic and non-genetic material.

SUMMARY

Two kinds of tetrad, ordered and unordered, are distinguished by the constant spacial position of spores in the former. The interpretation of various spacial patterns in ordered tetrads is described.

Both ordered and unordered tetrads are classified on the basis of the relations of pairs of loci. These relations can be two-type parental (P), two-type non-parental (N), or four-type (F). The relations of various patterns of single and multiple cross-overs to the classes of tetrad are given in tabular form.

On the assumption of no chiasma or chromatid interference, the addition law for F (= 2nd division segregation) frequencies is shown to be $F_{1,2} = F_1 + F_2 - 3F_1F_2/2$; and the relation of recombinant frequencies to F frequencies, $R = 1/2 - [(1 - 3F/2)^{2/3}]/2$.

Evidence bearing on chromatid and chiasma interference is extracted from various recent publications and a convenient statistical method, devised by L. J. SAVAGE, is applied to estimating interference with data pooled from numerous crosses involving small numbers.

APPENDIX I

The type of tetrad with respect to x loci may be represented by a $n \times 4$ rectangular array of 0's and 1's with exactly two zero's in every column. For example:

$$\begin{array}{cccccc} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array}.$$

Since the two zero's in each column can be chosen in 6 different ways, the total number of types possible, T(x), is, $T(x) = 6^x$.

Ordered tetrads

In an ordered tetrad notice is taken of the way in which cells lie as regards top and bottom two in a linear tetrad of four cells but no notice is taken of which pair are top and which are bottom, or of which single cell is top or bottom in any pair. Two types of tetrad of which one can be obtained from the other by interchanging individual cells of a pair, or by interchanging top with bottom pairs of cells, belong, so to speak, to the same species. The problem is to evaluate the number of species S(x).

The species may be conveniently divided into three genera thus:

Name	Description	No. of types per species in this genus	No. of species in the genus
G_1	w, x; y, z	8	$G_1(x)$
G_2	w, w; x, x	2	$G_2(x)$
G_3	w, x; w, x	4	$G_3(x)$

It is clear that $S(x) = G_1(x) + G_2(x) + G_3(x)$

$$T(x) = 6^x = 8G_1(x) + 2G_2(x) + 4G_3(x).$$

By the same argument which led to $T(x)$, there are 2^x types in G_2 and $2 \times 2^x = 2^{x+1}$ types in G_3, therefore $G_3(x) = (1/2)\ 2^x = 2^{x-1}$

$$G_3(x) = (1/4)\ 2^{x+1} = 2^{x-1}.$$

$$T(x) = 6^x = 8G_1(x) + 2^x + 2^{x-1} = 8G_1(x) + 3 \times 2^x$$
$$G_1(x) = (6^x - 3 \times 2^x)/8$$
$$S(x) = (6^x - 3 \times 2^x + 2^{x+2} + 2^{x+3})/8 = (6^x + 5 \times 2^x)/8$$

This formula was obtained by WHITEHOUSE (1942) empirically but was misprinted.

Unordered tetrads (complete classification)

In an unordered tetrad no notice is taken of the way in which cells lie as regards top two and bottom two and the genus may be divided into two species thus:

Name	Description	No. of types per species in this genus	No. of species in the genus
K_1	w, x, y, z	24	$K_1(x)$
K_2	w, w, x, x,	6	$K_2(x)$

$$S^u(x) = K_1(x) + K_2(x)$$
$$T(x) = 24K_1(x) + 6K_2(x).$$

By the previous argument there are 2^{x-1} types in K_2, therefore,

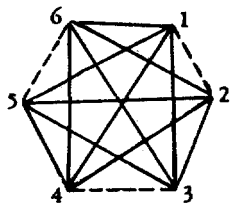
$$T(x) = 6^x = 24K_1(x) + 3 \times 2^x$$
$$K_1(x) = (6^x - 3 \times 2^x)/24 = (G^{x-1} - 2^{x-1})/4$$
$$S^u(x) = (6^{x-1} - 2^{x-1} + 4 \times 2^{x-1})/4 = (6^{x-1} + 3 \times 2^{x-1})/4.$$

Unordered tetrads (P, N, and F basis)

Any two columns (or loci), say c and d, of the $n \times 4$ matrix representing a tetrad must be in one of the three reflexive and mutually exclusive relations P, N, or F to each other: cPd, if c and d agree in every row; cNd, if c and d disagree in every row; cFd if c and d agree in two rows only.

Two tetrads are said to be of the same family if every pair of loci in one is in the same relation as the corresponding pair in the other. The problem is to compute the number $Q(x)$ of families when there are x loci. The computation of $Q(x)$ is based on certain considerations about the relations P, N, and F.

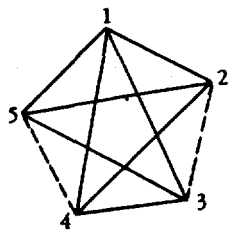
In the first place, if cPd, then to any other locus, say e, c and d must both bear the same relation. Thus P gathers the loci into clusters. What patterns are possible among clusters? Since there are but six different patterns of zeros and ones available, there can be at most six different clusters. If there are actually six they must be represented by some permutation of the pattern below, where solid lines denote F and dotted ones N.



15

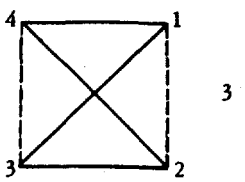
If six clusters are determined, as many possibilities may be formed from them as there are ways to arrange six things in pairs, or $6!/(2!)^3 3! = 15$. The number 15 has, for reference been written beside the figure.

Similarly five different clusters must be in a permutation of the pattern:

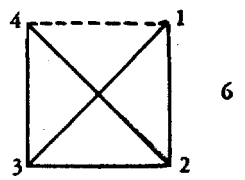


15

For four clusters two different sorts of patterns are possible:

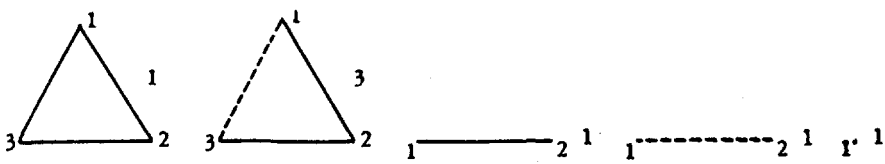


3



6

For three, two, and one the possibilities are these:



It is not a difficult combinatorial problem to see that the number of ways $P(k, x)$ in which x loci can be divided into k clusters is given by

$$k!\,P(k, x) = \sum_{0}^{k} (-1)^{k-j} \binom{k}{j} j^x \qquad (1)$$

If $P(k, x)$ were to be seriously studied for large x, the formula (1) would doubtless be useful. The following recursion formula would also be useful and it provides a convenient basis analogous to Pascal's triangle for computing $P(k, x)$ for small values of k and x.

$$P(k, x) = kP(k, x-1) + P(k-1, n-1)$$

$$Q(x) = P(1, x) + 2P(2, x) + 4P(3, x) + 9P(4, x) + 15P(5, x) + 15P(6, x).$$

The possible number of various kinds of tetrad and of random spores is given for values of x up to five in table 2.

TABLE 2

Numbers of possible tetrads and random spores when different numbers of loci are involved.

Number of loci	1	2	3	4	5
Total kinds of tetrad	6	36	216	1,296	7,776
Ordered tetrads	2	7	32	172	4,860
Unordered tetrads (complete)	1	3	12	60	336
Unordered tetrads (P, N, F, basis)	1	3	11	48	236
Random spores (2^x)	2	4	8	16	32

APPENDIX II

The statistical problem here is this. Several double dichotomies are observed, each constituted by a small number of observations. A statistical test is sought for the hypothesis that each is positively correlated.

Tradition would suggest that the hypothesis be tested by computing χ^2 for each double dichotomy, adding the values of χ^2 and appealing to the χ^2 test. This solution is rather unsatisfactory in two respects both of which may be met by a variation of it.

First, since each double dichotomy is small, the expectations of the individual hypotheses will, under the null hypothesis, each be a little less than one, so that the χ^2 test based on this sum will be systematically inaccurate. Since $\chi^2 = t^2(n-1)/n$, it may be noted that, as was asserted earlier $E(\chi^2) = (1-1/n)$ $E(t^2) = (1-1/n) < 1$.

Second, even if the double dichotomies were not small, adding the individual values of χ^2 leads to a test which is, at least to all appearances, less sensitive than a test based on the sums of χ itself would be.

To construct the test we propose, let a double dichotomy be represented by the familiar table.

a	b	y
c	d	z
w	x	n

where $w = a + c$, $n = w + x$, etc., and let $\Delta = ad - bc = an - yw$, and compute its expected value and variance of Δ under the null hypothesis thus:

$$E(\Delta) = nE(a) - yw = 0$$
$$V(\Delta) = n^2V(a) = wxyz/n - 1,$$

the expected value and variance of a is given in any account of the hypergeometric distribution. The quantity $t = \Delta\sqrt{(n-1)/wxyz}$ therefore has, under the null hypothesis, zero expectation and unit variance, provided that $wxyz > 0$. The sum of several values of it should therefore, under the null hypothesis, be approximately normally distributed about zero, with variance equal to the numbers of terms in the sum. Under the envisaged alternatives it will tend to be positive.

LITERATURE CITED

BELLING, J., 1933 Crossing over and gene rearrangement in flowering plants. Genetics **18:** 388–413.

BUSS, H. R., 1944 The genetics of methionineless mutants of *Neurospora crassa*. M. A. Thesis, Stanford University, 60 pp.

CHARLES, D. R., 1938 The spatial distribution of cross overs in X-chromosome tetrads of *Drosophila melanogaster*. J. Genet. **36:** 103–126.

FINCHAM, J. R. S., 1951 A comparative genetic study of the mating-type chromosomes of two species of *Neurospora*. J. Genet. **50:** 221–229.

HALDANE, J. B. S., 1919 The combination of linkage values, and the calculation of distances between loci of linked factors. J. Genet. **8**: 299–309.

HOULAHAN, M. B., G. W. BEADLE and H. G. CALHOUN, 1949 Linkage studies with biochemical mutants of *Neurospora crassa*. Genetics **34**: 493–507.

JANSSENS, F. A., 1909 La théorie de la chiasmatype. La Cellule **25**: 389–411.

KNAPP, E., 1937 Crossing-over und Chromosomenreduktion. Zeit. Ind. Abs. Ver. **73**: 409–418.

LINDEGREN, C. C., and G. LINDEGREN, 1942 Locally specific patterns of chromatid and chromosome interference in *Neurospora*. Genetics **27**: 1–24.

LINDEGREN, C. C., 1949 Chromosome maps of Saccharomyces. Hereditas Suppl. Vol., 338–355.

LUDWIG, W., 1938 Faktorenkoppelung und Faktorenaustausch bei normalem und aberrantem Chromosomenbestand. Georg Thieme, Leipzig. pp. 245, 74abb.

MATHER, K., 1935 Reductional and equational separation of the chromosomes in bivalents and multivalents. J. Genet. **30**: 53–78.

MATHER, K., and G. H. BEALE, 1942 The calculation and precision of linkage values from tetrad analysis. J. Genet. **43**: 1–30.

MATSUURA, H., and T. HAGA, 1942 Chromosome studies on *Trillium kamtschaticum* Pall. X. On the origin of the chiasma. Cytologia **12**: 397–417.

MCCLINTOCK, B., 1945 Preliminary observations of the chromosomes of *Neurospora crassa*. Amer. J. Bot. **32**: 671–678.

MOEWUS, F., 1940 Die Analyse von 42 erblichen Eigenschaften der *Chlamydomonas eugametos*-gruppe. III die 10 Koppelungsgruppen. Zeit. Ind. Abs. Ver. **78**: 501–522.
1949 Zur Genetik und Physiologie der Kern un Zellteilung. II Uber den Synchronismus der Kernteilungen bei *Protosiphon botryoides*. Beit. Biol. Pflanzen. **28**: 36–63.

PAPAZIAN, H. P., 1951 The incompatibility factors and a related gene in *Schizophyllum commune*. Genetics **36**: 441–459.

PERKINS, D. D., 1949 Biochemical mutants in the smut fungus *Ustilago maydis*. Genetics **34**: 607–626.

SHENG, T. C., 1951 A gene that causes natural death in *Neurospora crassa*. Genetics **36**: 199–212.

WEINSTEIN, A., 1936 The theory of multiple-strand crossing-over. Genetics **21**: 155–199.

WETTSTEIN, F. V., 1924 Morphologie und Physiologie des Formwechsels der Moose auf genetischer Grundlage. I. Zeit. Ind. Abs. Ver. **33**: 1–236.

WHITEHOUSE, H. L. K., 1942 Crossing-over in Neurospora. New Phytol. **41**: 23–62.
1949 Multiple-allelomorph heterothallism in the fungi. New Phytol. **48**: 212–244.
1950 Mapping chromosome centromeres by the analysis of unordered tetrads. Nature **165**: 893.

WINGE, O., 1947 The segregation in the ascus of *Saccharomycoides ludwigii*. C. R. Lab. Carlsberg (phys.) **24**: 223–226.

WÜLKER, H., 1935 Untersuchungen über Tetradenaufspaltung bei *Neurospora sitophila* Shear et Dodge. Zeit. Ind. Abs. Ver. **69**: 210–248.