

# THE RELATIONSHIP BETWEEN DIPLOID AND TETRAPLOID RECOMBINATION FREQUENCIES

J. A. SVED

*Department of Genetics, University of Adelaide, South Australia*

Received 2.iv.64

## 1. INTRODUCTION

THE recombination frequencies in similar segments of diploids and autotetraploids have been compared in several experiments (de Winton and Haldane, 1931; Sansome, 1933; Oram, 1959; Welch, 1962). However, there appears to have been little discussion on the expected relationship between these two variables. The hypothesis of equality for two such frequencies can only be given as a first approximation, since multiple crossing-over may be expected to affect the two in different ways. In general it appears that multiple crossing-over may lead to a higher tetraploid recombination frequency than the corresponding frequency in diploids, even though the mean frequency of crossing-over may not be increased.

In this paper the relationship between the number of crossovers and the frequency of recombination in a particular region will be given for a number of limiting models for diploids and tetraploids. The distribution of the number of crossovers will then be specified, allowing the relationship between the two recombination frequencies to be derived. An upper bound will be given for the amount by which the tetraploid recombination frequency could be expected in practice to exceed the corresponding diploid frequency, when the frequency of crossing-over is the same in the two cases.

## 2. THE RELATIONSHIP BETWEEN THE RECOMBINATION FREQUENCY AND THE NUMBER OF CROSSEVERS

The recombination frequency between two loci may be expressed in the following form:—

$$y = a_0q_0 + a_1q_1 + a_2q_2 + \dots$$

where  $q_i$  are the frequency with which exactly  $i$  crossovers are found between the two loci, and  $a_i$  are the mean recombination frequency given by  $i$  crossovers. (It will be convenient to use the term "crossover" rather than "chiasma" in most places throughout this paper. Each chiasma is assumed to be the cytological manifestation of a crossover, and to give rise to an exchange of material between two strands.) The  $a_i$ 's of this relationship are determined solely by the configuration of these crossovers. In this section the evaluation of the  $a_i$ 's will be attempted for diploids and tetraploids.

### (1) *Diploids*

Since there are only two chromosomes involved for this case, the  $a_i$ 's depend only on the amount of chromatid interference. If chromatid

interference is absent, as will be assumed throughout this paper, the relationship is a simple one (Mather, 1938),

$$a_0 = 0$$

$$a_i = \frac{1}{2} \text{ for } i = 1, 2, 3, \dots$$

### (ii) Tetraploids

The analysis of gametic data in tetraploid organisms is complicated by the fact that there are two homologous chromatids in each gamete, and considering two linked loci, there are eleven modes of gamete formation (Fisher, 1947). For evaluation of the recombination fraction, however, strands are considered separately, and for two loci on a single strand there are only two modes of formation, viz. recombinant and non-recombinant. Whereas the position of the centromere will have considerable effect on the frequencies of the modes of gamete formation, it is not of primary importance when strands are considered separately, and under the assumptions to be considered it may be neglected altogether. Thus calculation of the recombination frequency, although it yields all the relevant information for two loci in diploids, yields only portion of the information available from the diploid gametes of tetraploids.

Besides the assumption of absence of chromatid interference, additional assumptions about chromosome pairing must be considered for deriving the  $a_i$  for tetraploids. Four models will now be considered.

*Model 1—bivalent formation.* If there is no quadrivalent formation, no change is expected from the relationship derived for diploids, viz.

$$a_0 = 0 \quad a_i = \frac{1}{2} \quad i = 1, 2, 3, \dots$$

For this relationship the bivalents have been considered separately, and  $i$  refers to the number of crossovers per bivalent. The formation of chiasmata in one bivalent may not be independent of their formation in the other, since bivalent competition may occur (Mather, 1936a). However, this phenomenon cannot affect the relationship for a single bivalent, and for the moment can be ignored.

*Model 2—random change of partner.* All eight strands must be considered together in this model and  $i$  refers to the total number of crossovers between the two loci per quadrivalent. Strands are associated pairwise at all points, and any strand is equally likely to be paired with any other strand at a given point. A typical configuration may be depicted as in fig. 1. A genotype with four distinctive A and B alleles is used in formulating expectations in this and the following tetraploid models. However, as shown by Fisher (1947), the recombination frequency can be estimated even if only two alleles are available at each locus, and all conclusions derived using the tetragenic models still apply to the recombination frequency estimated from digenic material. The assumptions of random changes of partner and

no chromatid interference taken jointly mean that each of the twenty-four possible non-sister crossovers is equally likely at any point.

Obviously  $a_0$  must equal 0, and  $a_1 = 1/4$ . By direct enumeration, *i.e.* by writing out all configurations possible with two crossovers,  $a_2$  can be shown to be  $5/12$ , and similarly other  $a_i$  could be found. However, a recurrence relation can readily be found for the  $a_i$  by considering the fate of a single strand.

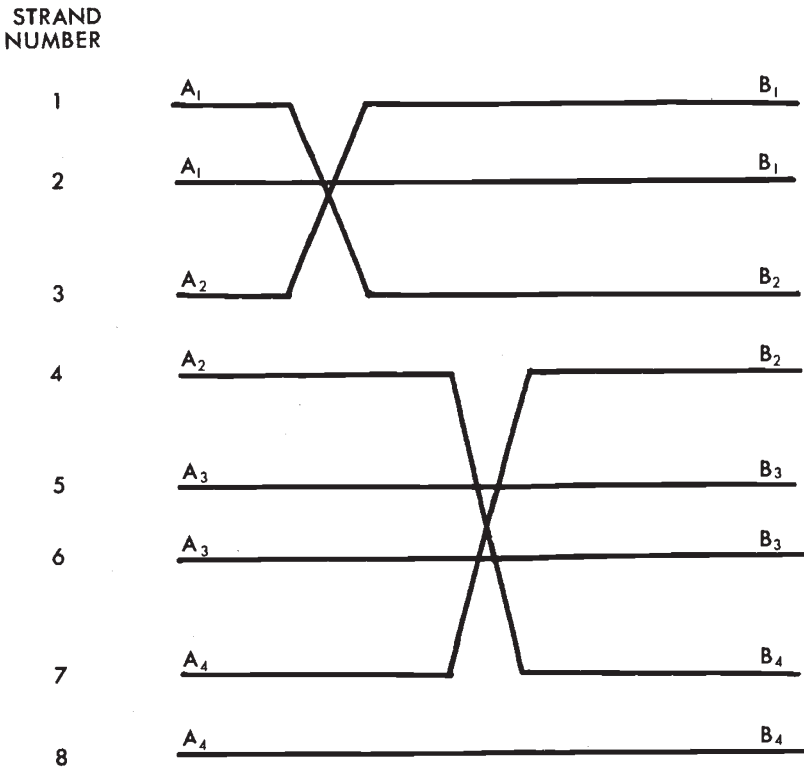


FIG. 1.—Diagrammatic representation of random change of partner model showing a typical configuration with two crossovers.

Regarding only one of the A alleles, *e.g.*  $A_1$  on strand 1, consider the crossover configuration after  $i$  crossovers between the two loci, and the effect of adding another crossover adjacent to these. The probability that  $A_1$  is connected to either  $B_2$ ,  $B_3$  or  $B_4$  after  $i$  crossovers between the two loci is by definition  $a_i$ . Then two of the twenty-four crossovers which could be added adjacent to the first  $i$  crossovers will alter this situation, *i.e.* restore the connection between  $A_1$  and  $B_1$ , and the other twenty-two will not. If, for example,  $A_1$  is connected to  $B_2$  on strand 3, then the two crossovers leading to non-recombinant strands would be those involving strands 1 and 3, or strands 2 and 3. Similarly, if  $A_1$  were connected to  $B_1$  after  $i$  crossovers, with probability  $1 - a_i$ , then six of the twenty-four possible crossovers would lead to a

recombinant strand. Thus the probability that  $A_1$  is connected to  $B_2, B_3$  or  $B_4$  after  $i + 1$  crossovers is

$$a_{i+1} = 22/24 \cdot a_i + 6/24(1 - a_i)$$

$$\therefore a_{i+1} = 1/4 + 2/3a_i.$$

Thus  $a_i$  must be of the form  $e + m\left(\frac{2}{3}\right)^i$ , and since  $a_1 = 1/4$  and  $a_2 = 5/12$ , we can substitute to find  $l$  and  $m$  giving  $l = -m = 3/4$

$$\therefore a_i = 3/4 \left[ 1 - \left(\frac{2}{3}\right)^i \right].$$

One result from this calculation, which does not appear to have been reported previously, is that under this model the maximum recombination frequency in tetraploids is  $3/4$ . This seems intuitively

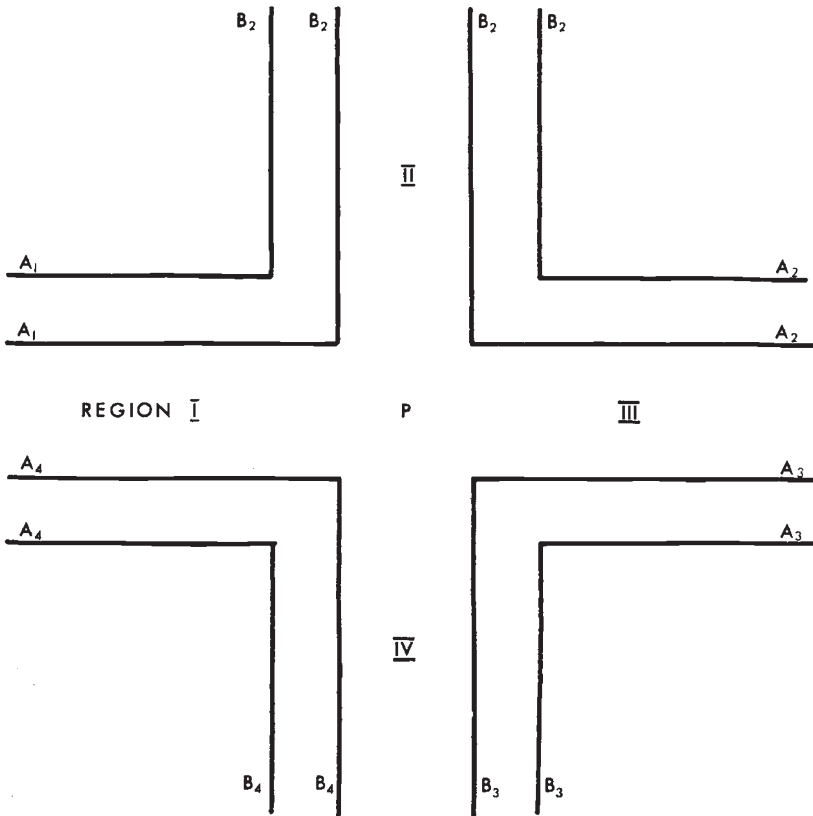


FIG. 2.—Diagrammatic representation of quadrivalent with one change of partner per chromosome.

reasonable since after an infinite number of crossovers the A's and B's could be regarded as oriented at random, giving a one in four chance that any  $A_j$  would be connected to  $B_j$ .

*Model 3—one change of partner, no interference.* The model of random

changes of partner, although mathematically tractable, is unlikely to be attained in practice, and the more realistic assumption of one change of partner per chromosome is considered here. This model, which is depicted in fig. 2, assumes that chromosomes are associated in pairs at all points, with one point of partner change occurring at a homologous point on each of the four chromosomes. Cytological observation suggests that this type of pairing may occur frequently.

It may be seen from fig. 2 that regions I and III and regions II and IV are equivalent. The ratio of region I to region II, however, need not necessarily be constant, and in the following it will be assumed that the point of partner change may lie at any point between A and B with equal probability.

Let the ratio AP/AB be  $x$ . Then the probability that a crossover will lie in region I is  $x/2$ , similarly for region III, and the corresponding probabilities for regions II and IV are  $\frac{1}{2}(1-x)$ . With a complete noninterference model, the distribution of  $i$  crossovers between regions I, II, III and IV is given by the terms in the multinomial

$$\left(\frac{x}{2} + \frac{1-x}{2} + \frac{x}{2} + \frac{1-x}{2}\right)^i.$$

If all crossovers lie in the same region, say region I, then since there is no chromatid interference the recombination frequency between A and B will be  $1/4$ . If crossovers lie in two adjacent regions, however, say regions I and II, one-quarter of the  $A_1B_1$  strands will remain unbroken, and likewise one-half of the  $A_2B_2$  and  $A_4B_4$  and all the  $A_3B_3$  strands will be unbroken, giving the overall recombination frequency as

$$\frac{1}{4}\left(\frac{3}{4} + \frac{1}{2} + 0 + \frac{1}{2}\right) = \frac{7}{16}.$$

Other distribution types give recombination frequencies as follows: regions I and III,  $1/2$ ; regions I, II and III,  $5/8$ ; and regions I, II, III and IV,  $3/4$ .

The probabilities of the different distribution types may now be evaluated from the multinomial, giving the recombination frequency as

$$\begin{aligned} & \frac{1}{4} \left[ 2 \left(\frac{x}{2}\right)^i + 2 \left(\frac{1-x}{2}\right)^i \right] + \frac{7}{16} \left[ 4 \left(\frac{1}{2}\right)^i - 4 \left(\frac{x}{2}\right)^i - 4 \left(\frac{1-x}{2}\right)^i \right] \\ & + \frac{1}{2} \left[ x^i - 2 \left(\frac{x}{2}\right)^i + (1-x)^i - 2 \left(\frac{1-x}{2}\right)^i \right] \\ & + \frac{5}{8} \left[ 2 \left(\frac{2-x}{2}\right)^i + 2 \left(\frac{1+x}{2}\right)^i - 8 \left(\frac{1}{2}\right)^i - 2x^i - 2(1-x)^i + 6 \left(\frac{x}{2}\right)^i \right. \\ & \qquad \qquad \qquad \left. + 6 \left(\frac{1-x}{2}\right)^i \right] \\ & + \frac{3}{4} \left[ 1 - 2 \left(\frac{2-x}{2}\right)^i - 2 \left(\frac{1+x}{2}\right)^i + 4 \left(\frac{1}{2}\right)^i + x^i + (1-x)^i \right. \\ & \qquad \qquad \qquad \left. - 2 \left(\frac{x}{2}\right)^i - 2 \left(\frac{1-x}{2}\right)^i \right]. \end{aligned}$$

Since  $x$  is assumed to be distributed with frequency function  $dx$  in the range (0, 1), the mean recombination frequency for  $i$  crossovers is obtained by integrating the above expression for  $x$  in the range (0, 1), giving on simplification

$$a_i = \frac{3}{4} - \left(\frac{1}{2}\right)^{i+2} - \frac{1}{i+1} + \left(\frac{1}{2}\right)^{i+1} \frac{1}{i+1}.$$

*Model 4—one partner exchange, crossover repulsion.* The same conditions as in model 3 apply, except that crossovers instead of forming independently are now assumed to repel each other. Thus, if two crossovers are formed between the two loci, they will repel each other into opposite arms of the quadrivalent. Similarly three crossovers will lie on three different arms, and four on four different arms, while the positioning of any subsequent crossovers beyond four is arbitrary. No increase in the frequency of crossovers is postulated, but merely a spreading out in the region AB. Under this model

$$a_0 = 0, \quad a_1 = 1/4, \quad a_2 = 1/2, \quad a_3 = 5/8, \\ a_i = 3/4, \quad i = 4, 5, 6, \dots$$

The importance of this model is that it supplies an upper bound to the frequency of recombination which may occur for a given number of crossovers. The maximum recombination frequency of 3/4 is only attained in the limit with an infinite number of crossovers under models 2 and 3, whereas under this model it is reached with four or more crossovers. In addition it may readily be seen that, in the absence of chromatid interference, for neither one, two or three crossovers can a greater frequency of recombination be given than that predicted by this model.

### 3. THE UPPER LIMIT OF RECOMBINATION

In the absence of chromatid interference, it is evident that with one or more changes of partner between two loci there is an upper limit of 75 per cent. for the recombination frequency. The case where a quadrivalent is formed with all points of partner change lying outside the interval between the two loci must be regarded with respect to the recombination frequency between these loci as equivalent to bivalent formation. (However, when considering the fate of both chromatids in a gamete the formation of a partner change outside the two loci cannot be ignored.) In general, where a point of partner change lies between two loci in a fraction  $p$  of cases, the upper limit for the recombination frequency becomes

$$p(0.75) + (1-p)(0.50) = 0.50 + 0.25p.$$

By the argument advanced in the previous section, it appears intuitively correct to say that the upper limit of recombination in an  $x$ -ploid

organism is  $\frac{x-1}{x}$ . The attainment of this recombination fraction, however, depends on the occurrence of a sufficiently high frequency of partner changes. With hexaploids for example, the limit  $5/6$  is only reached as the number of partner changes becomes infinite, the approach to the limit being oscillatory, and certain combinations of partner changes may cause the limit to be exceeded slightly.

#### 4. THE RECOMBINATION FREQUENCY BETWEEN UNLINKED GENES

The recombination frequency when defined in the usual way has no meaning when applied to two unlinked genes in a tetraploid. However, it may be shown that the amount of recombination given by two unlinked genes is equivalent to that given by two genes linked with a recombination frequency of  $3/4$ . Consider the gametic array given by two linked genes under the model of random partner exchange (model 2). The properties of the transition matrix defined by this model are such (see *e.g.* Moran, 1962) that after a sufficient number of crossovers the genes at the two loci will be combined essentially at random, which is equivalent to the array given by unlinked genes. The limiting recombination frequency for the model of random partner exchange has been shown to be  $3/4$ , thus demonstrating the above assertion.

In practical terms, however, this result cannot be taken to mean that the amount of recombination between unlinked genes is greater in the tetraploid than in the diploid. In this context it is perhaps more relevant to consider both alleles in the gamete of the tetraploid. For the tetraploid, as well as the diploid, two randomly chosen unlinked genes in the zygote have one-quarter chance of both being present in the gamete, one-half chance of one being present and the other absent and one-quarter chance of both being absent. Double reduction will, in fact, reduce the proportion of the recombinant classes in the tetraploid to a little below one-half.

#### 5. THE RELATIONSHIP BETWEEN DIPLOID AND TETRAPLOID RECOMBINATION FREQUENCIES

In order to specify the recombination frequency for either diploids or tetraploids, it now becomes necessary to give values for  $q_i$  in the relation

$$y = a_0q_0 + a_1q_1 + \dots + a_iq_i + \dots$$

Since no realistic crossover distributional theory exists, it is necessary as a first approximation to calculate  $y$  under the assumption of no chromosome interference, and ultimately to consider what effect chromosome interference will have on this calculation.

The assumption of no chromosome interference is equivalent to

assuming a Poisson distribution of crossovers. If the mean number of crossovers is  $m$ , we have

$$q_i = \frac{e^{-m} m^i}{i!}.$$

The recombination fraction for diploids,  $y_D$ , is equal to  $\frac{1}{2}q_1 + \frac{1}{2}q_2 + \dots = \frac{1}{2}(1 - q_0)$ , giving the well-known relation, (cf. Haldane, 1919)

$$y_D = \frac{1}{2}(1 - e^{-m}) \dots \quad (1)$$

The relationship between  $y_D$  and the tetraploid recombination fraction  $y_T$  is dependent first on the relative frequencies of crossovers between the two loci in the diploid and tetraploid, and secondly on which of the four tetraploid models is used. The relationships in this section will be derived assuming that the mean frequency of crossing-over per strand is the same in the diploid and tetraploid, *i.e.* that the mean number of crossovers in the tetraploid is twice that in the diploid. This particular choice will be discussed briefly in the following section.

*Model 1.* With no assumptions about the crossover distribution necessary in this case, we have

$$y_T = y_D.$$

*Model 2.* The value of  $q_i$  for this case may be given as  $\frac{e^{-m'} (m')^i}{i!}$ , where  $m'$  is the mean number of crossovers for a quadrivalent. Since we wish to compare the recombination frequencies for diploid and tetraploid when the mean frequencies of crossing-over are the same, we must place  $m' = 2m$ .

Then  $y_T = a_0 q_0 + a_1 q_1 + \dots$

$$\begin{aligned} &= \frac{3}{4} \sum_{i=0}^{\infty} \left[ 1 - \left( \frac{2}{3} \right)^i \right] \cdot \frac{e^{-2m} (2m)^i}{i!} \\ &= \frac{3}{4} \left( 1 - e^{-\frac{2m}{3}} \right). \end{aligned}$$

Then substituting for  $m$  from equation (1) we get

$$y_T = \frac{3}{4} \left[ 1 - (1 - 2y_D)^{\frac{2}{3}} \right]. \quad (2)$$

This is given as curve (2) in fig. 3.

*Models 3 and 4.* Once again for both models

$$q_i = \frac{e^{-2m} (2m)^i}{i!}.$$

Substituting the values previously found for  $a_i$  gives for model 3

$$y_T = \frac{3}{4} - \frac{1}{4} e^{-m} - \frac{1}{2m} [1 - e^{-m}].$$



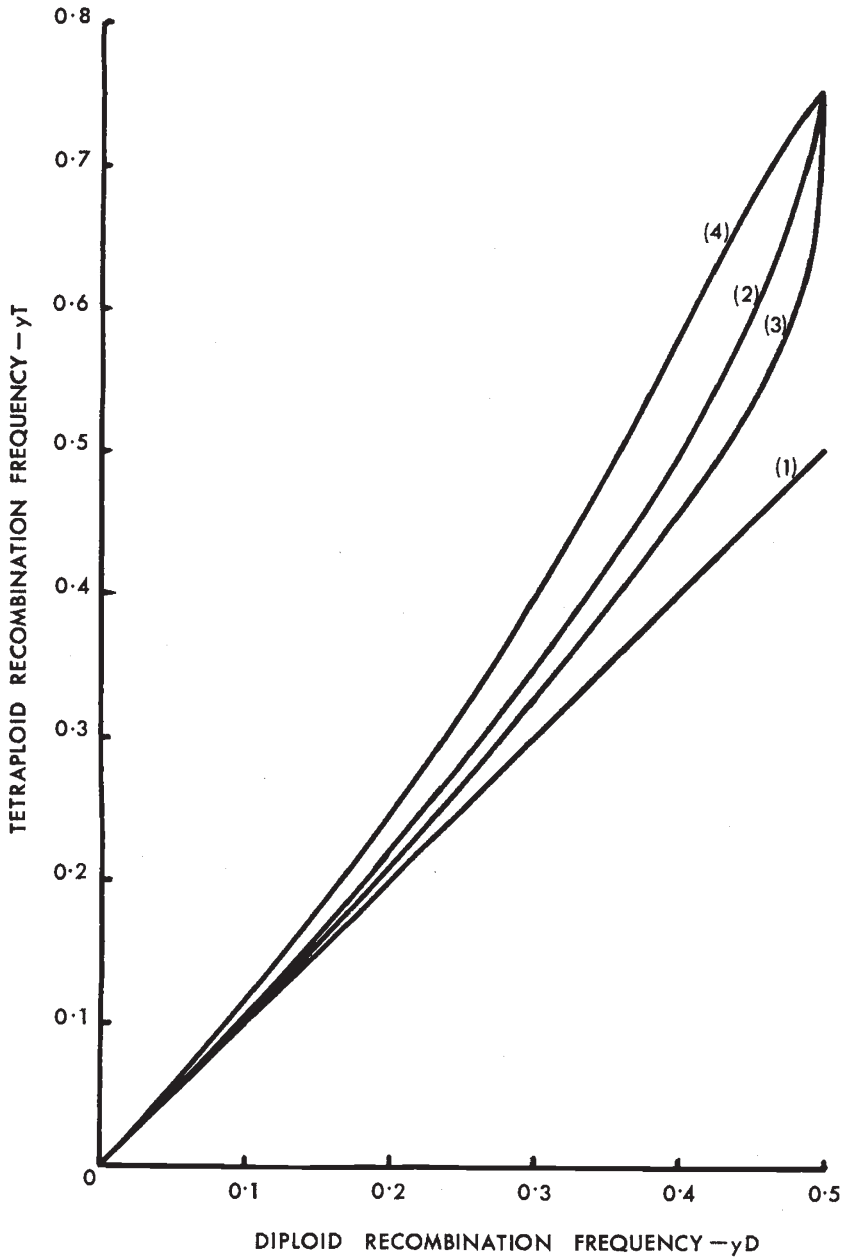


FIG. 3.—The relationships between diploid and tetraploid recombination frequencies under four tetraploid models.

Substituting for  $m$  from equation (1)

$$y_T = \frac{y_D}{\log(1-2y_D)} + \frac{y_D}{2} + \frac{1}{2}. \quad (3)$$

Similarly for model 4 we have

$$y_T = \frac{3}{4} - e^{-2m} \left[ \frac{3}{4} + m + \frac{1}{2} m^2 + \frac{1}{6} m^3 \right] \quad (4)$$

which together with equation (1) forms the parametric equations for the relation between  $y_T$  and  $y_D$  graphed as curve (4).

## 6. DISCUSSION

Although the above relationships between  $y_T$  and  $y_D$  have been derived on the assumption of equal frequencies of crossing-over in the tetraploid and diploid, they could equally well be calculated assuming different rates of crossing-over in the two cases. The relationships given are not those necessarily expected in practice between the two frequencies, but those (for the particular models) expected on the hypothesis that the process of crossing-over is not affected by a doubling of the chromosome set, including such events as changes of partner, etc. Since the frequency of crossing-over rather than the frequency of recombination is the fundamental biological quantity, this appears to be a more appropriate hypothesis to take than the hypothesis of equality of the two recombination frequencies.

It may be seen from the four curves in fig. 3 that considerably different relationships between  $y_T$  and  $y_D$  are expected depending upon which of the four tetraploid models is used. In view of the fact that none of the models can be expected to be exactly applicable, it does not seem possible to give a single relationship which will apply in practice. At best, upper and lower bounds for the value of  $y_T$  for a given  $y_D$  can be calculated. It will be argued below that under conditions expected to apply in practice, curves (4) and (1) represent such upper and lower bounds.

As pointed out previously the value  $y_T$  given by model 4 represents an upper bound to the amount of recombination which may be observed for a given number of crossovers. The curve (4) is therefore an upper bound for  $y_T$  given  $y_D$ , and assuming no chromosome or chromatid interference. Since the positive value of the difference  $y_T - y_D$  is attributable to multiple crossing-over, it is evident that positive chromosome interference must cause this difference to be reduced. The effect of chromatid interference is more difficult to assess, but it appears that a small amount of positive chromatid interference would tend to lower the difference  $y_T - y_D$ , and negative interference to raise it. Thus it appears that in the absence of negative chromosome or negative chromatid interference, neither of which is likely in practice, the curve (4) represents an upper bound for  $y_T$  for

a given  $y_D$  when the crossover frequencies are the same in the corresponding diploid and tetraploid segments. The value of  $y_D$  could only exceed the value  $y_T$  when there was bivalent competition, the expected excess even for a strong competition being trivial, so that curve (1) can be given as a lower bound for  $y_T$  for a given  $y_D$ .

One use of the upper bound may be illustrated as follows. When comparing data on the recombination frequencies for two loci in diploids and tetraploids, two questions may be asked:—

1. Whether the two recombination frequencies differ significantly, and
2. Whether, if the tetraploid recombination frequency is greater than the diploid frequency, the difference falls significantly outside the bounds given by equation (4), *i.e.* whether the increased recombination frequency necessarily reflects a rise in the frequency of crossing-over.

The only data which appear suitable for these comparisons are those of Oram (1959). Those studies in which recombination frequencies are calculated from first backcross data could not show increases in the tetraploid due to change of partner and multiple crossing-over, since the recombination frequency is calculated on the assumption of single crossing-over. As discussed by Mather (1936*b*) these estimates are unlikely to be accurate at distances of over fifteen map units.

Considering the sugary and glossy loci in maize, Oram finds:—

$$y_D = 0.282 \pm 0.009$$

$$y_T = 0.556 \pm 0.062.$$

The difference  $y_T - y_D$  may be shown to be significant at the one per cent. level of significance. The second test of significance may be made by plotting the point (0.282, 0.556), and its associated 95 per cent. confidence limits in fig. 2. All points within the confidence interval are found to lie to the left of the curve given by equation (4). Thus an excess in  $y_T$  of this order could not be attributed to the effect discussed in this paper, but could, however, reflect a rise in the frequency of crossing-over in the tetraploid.

## 7. SUMMARY

It is shown that multiple crossing-over may be expected to affect the diploid and tetraploid recombination frequencies in different ways. As a result the tetraploid recombination frequency for a particular region may exceed the corresponding diploid frequency even when the mean frequency of crossing-over is the same in the two cases. An upper bound is given for the amount of such an excess which may be expected in practice.

In the absence of chromatid interference, the upper limit of recombination in tetraploids is shown to be  $\frac{3}{4}$ . In general, in an  $x$ -ploid species, provided the number of changes of partner is sufficiently high, the upper limit of recombination is  $\frac{x-1}{x}$ . It is shown that the amount of recombination given by unlinked genes in tetraploids is equivalent to that given by two genes linked with a recombination frequency of  $\frac{3}{4}$ .

A discussion is included of the interpretation of data from experiments in which the diploid and tetraploid recombination frequencies are compared. The data from such an experiment of Oram (1959) are used in illustrating this analysis.

*Acknowledgments.*—I am indebted to Professor J. H. Bennett and Mr I. R. Franklin for their assistance in discussions; to Dr R. N. Oram for permission to quote his unpublished data, and to C.S.I.R.O. for their award of a studentship during the tenure of which this work was done.

## 8. REFERENCES

- DE WINTON, D., AND HALDANE, J. B. S. 1931. Linkage in the tetraploid *Primula sinensis*. *J. Genet.*, 24, 121-144.
- FISHER, R. A. 1947. The theory of linkage in polysomic inheritance. *Phil. Trans. Roy. Soc. B.*, 232, 55-87.
- HALDANE, J. B. S. 1919. The combination of linkage values, and the calculation of distance between the loci of linked factors. *J. Genet.*, 8, 299-309.
- MATHER, K. 1936a. Competition between bivalents during chiasma formation. *Proc. Roy. Soc. B.*, 120, 208-227.
- MATHER, K. 1936b. Segregation and linkage in autotetraploids. *J. Genet.*, 32, 287-314.
- MATHER, K. 1938. Crossing over. *Biol. Rev.*, 13, 252-292.
- MORAN, P. A. P. 1962. Polysomic inheritance and the theory of shuffling. *Sankhya, A.*, 24, 63-72.
- ORAM, R. N. 1959. Genetics of autotetraploid maize. *Ph.D. Thesis*, University of Adelaide.
- SANSOME, F. W. 1933. Chromatid segregation in *Solanum lycopersicum*. *J. Genet.*, 27, 105-132.
- WELCH, J. E. 1962. Linkage in autotetraploid maize. *Genetics*, 47, 367-396.