THE GENERAL THEORY OF MAPPING FUNCTIONS FOR RANDOM GENETIC RECOMBINATION

R. H. WALMSLEY

From the Physics Department, University of Pennsylvania, Philadelphia, Pennsylvania 19104

ABSTRACT The general theory of mapping functions for random genetic recombination is derived without reference to physical models. The results are of particular relevance to asymmetric recombination (preferential recombination of the alleles of either parent) and include the effects of lethal lesions. It is noted that the assumption of randomness determines the mapping functions uniquely so that they are independent of the specifications of any stochastic physical model of recombination. In particular, the functions derived from physical models by Haldane, Wu, Wood and Verhoef and de Haan are special cases of those derived here. A general principle for the testing of genetic linkage is formulated which is valid for both symmetric and asymmetric recombination.

INTRODUCTION

Observed linkage and recombination frequencies between genetic markers have given rise to more and less explicit models to represent the effects of the processes underlying genetic recombination. The two most commonly discussed are the break-reunion model and the copy-choice model, together with their several variations. In all cases, however, the quantitative task is the same: to account for the observed relative frequencies of linkage and recombination by deriving mapping functions which depend upon the physical distance between the markers.

Assuming randomness in the distribution of cross-over events, Haldane (1919) analyzed the simple break-reunion model and derived a mapping function for the recombination frequency, R, given by

$$R = \tfrac{1}{2}(1 - e^{-2\nu x}), \tag{1}$$

where x is the physical distance between genes and ν is the mean density of cross-over events. More sophisticated versions of equation 1 have been evolved (Bailey, 1961) to correct for interference phenomena observed in higher organisms. However, equation 1 gives $R \approx \nu x$, for $\nu x \ll 1$, and $R \approx 0.5$, for $\nu x \gg 1$, which are in accord with much of the data for eukaryotes and for bacteria as well (Jacob and

Wollman, 1961). Wu (1967) has applied Haldane's analysis to genetic recombination data in *Escherichia coli*. He estimates a value for the mean cross-over density of 0.1/min of map distance.

One of the intrinsic assumptions of the simple break-reunion model is that the two parental chromosomes enter the recombination process in an exactly reciprocal fashion. The asymptotic value of 0.5 for R at large values of x is a direct consequence of this assumption. When the correlation between markers has been lost due to the "noise" of the fluctuations in crossing-over, i.e. $\nu x \gg 1$, there is no bias toward either chromosome. Several workers have shown, however, that asymmetric recombination can result from matings of irradiated or nonisogenic bacteria. Using *E. coli* K-12 bacteria, Wood and Marcovich (1964) irradiated the recipients with X-rays and then mated them with normal donors. Genetic analysis of the recombinants showed that the frequency of occurrence of the male alleles of the unselected markers became increasingly biased toward the donor and approached 1.0 at large doses. This experiment highlights the fact that genetic analysis is performed on viable recombinants only. A general formalism for mapping functions should therefore incorporate the effects of lethal lesions. More recently, Verhoef and de Haan (1966) have reported genetic analysis results, from both isogenic and nonisogenic matings with *E. coli*, which yielded asymptotic values of R strongly biased toward the recipient. They found a value of 0.23 for one isogenic K-12 mating and a value of 0.019 for a mating in which a strain K-12 donor was crossed with a strain B recipient.

It is clear, then, that the intrinsically symmetric analysis of Haldane is too restricted. Wood (unpublished) has made an explicit analysis of the copy-choice model in which there are separate sets of "switching" events for the two parental chromosomes. As the genetic information for the recombinant chromosome is copied from either parental chromosome, there is a probability of being switched away to copy from the other. In general, these two sets of events may have different mean densities, with the result that the recombinant may contain more information, on the average, from one parent than from the other. Verhoef and de Haan (1966) have proposed a variation of the break-reunion model in which the breakage events are considered to be distributed randomly with a constant mean density and segments of the donor chromosome are randomly incorporated with a fixed probability. Here again there are two sets of random events, instead of one, which allows for the asymmetric incorporation of one or the other set of parental alleles.

The question naturally arises as to whether, by comparing the theoretical mapping functions to experimental data, it is possible to judge the correctness of the models from which they were derived. Furthermore, the asymmetric models discussed above contain two adjustable parameters instead of one. It is obvious that they can be made to give a better goodness-of-fit than the functions for the simple break-reunion model, which contains only one parameter. The question here is whether or not it is meaningful to propose more and more detailed models with in-

creasing numbers of adjustable parameters to give better fits to better data. It is the purpose of this paper to answer these questions and to put the subject of mapping functions on a broader theoretical footing. The discussion will deal with the general class of stochastic recombination processes, i.e., those which result in randomly distributed recombination events. We shall show that the mapping functions for all stochastic models are unique and can be derived without reference to any specific model; hence they are the most general ones possible. It can be shown, in fact, that the functions derived by Verhoef and de Haan are identical to those derived from the copy-choice model. The functions derived by Haldane are simply special cases of these. Finally, we shall include the effects of random lethal lesions on the chromosomes.

ASSOCIATION FREQUENCIES

In dealing with genetic recombination between two markers, a and b, it is necessary to measure the frequency of appearance in the progeny, relative to the number of mating pairs, of four classes: a^+b^+, a^+b^-, a^-b^+, and a^-b^-. Normally the sum of these relative frequencies is unity, but when lethal lesions are included this will not be true. Furthermore, in the common case of symmetric recombination it is expected that the relative frequencies of the a^+b^- and a^-b^+ classes will be the same, as will be the frequencies of the a^+b^+ and a^-b^- classes. In asymmetric recombination, however, this will not be true either. Therefore it will be necessary to derive mapping functions for all four classes: two recombination functions for the classes a^+b^- and a^-b^+, and two linkage functions for the classes a^+b^+ and a^-b^-. Having now four frequencies instead of the customary two, we shall use the term *association frequencies* to refer to them categorically. The corresponding probability functions (mapping functions) will be designated by the symbols $L_M{}^m$, $L_M{}^f$, $L_F{}^f$, and $L_F{}^m$, where m and f stand for male and female, respectively. The lower index indicates the parental allele of the selected marker, while the upper index indicates that of the unselected marker. For example, $L_M{}^f(x)$ is the probability of occurrence of the female allele of an unselected marker which is separated by a distance x from the selected marker, given the occurrence of the male allele of the selected marker.

Addition Theorems, Inversion Invariance, and Analysis of Multiple Marker Recombinant Classes

We shall represent the donor allelic forms of the genetic markers by the ordered sequence $a, b, c, \cdots$. The corresponding recipient allelic forms will be designated by bars over the letters, e.g., $\bar{a}$. The gene locations on the chromosome will be given, in the order of increasing x, by $x_a, x_b, x_c, \cdots$. Although these quantities are algebraic, it should be noted that the arguments of the association functions involve distances, that is, absolute magnitudes of coordinate differences. Using standard probability notation, we shall denote the probability of occurrence of, for example,

$a, \bar{b}, c, \cdots$, given the selected occurrence of $\bar{u}, \bar{v}, w, \cdots$, by $G(a, \bar{b}, c, \cdots \mid \bar{u}, \bar{v}, w, \cdots)$.

In an experiment involving N unselected markers, there are 2^N different recombinant classes whose probabilities of occurrence are to be calculated. Since the recombination events are assumed to be randomly distributed, the association frequencies between two sequential markers are independent of the events lying outside the interval between them. Therefore, a particular recombinant class with a given sequence of N alleles for the unselected markers has the probability of the associated Markov chain. This is given by the product of the corresponding association functions. For example, the probability of the class $\bar{a}b\bar{c}$, where $\bar{c}$ is the selected allele, is:

$$G(\bar{a}, b \mid \bar{c}) = L_F^m(x_c - x_b) \cdot L_M^f(x_b - x_a).$$

In the same fashion, if x and y are the lengths of an ordered sequence of two adjacent intervals, we obtain the addition theorems for the association functions from the sum rules for the corresponding Markov chains:

$$\begin{aligned}
L_M^m(x, y) &= L_M^m(x) \cdot L_M^m(y) + L_M^f(x) \cdot L_F^m(y) \\
L_M^f(x, y) &= L_M^m(x) \cdot L_M^f(y) + L_M^f(x) \cdot L_F^f(y) \\
L_F^f(x, y) &= L_F^f(x) \cdot L_F^f(y) + L_F^m(x) \cdot L_M^f(y) \\
L_F^m(x, y) &= L_F^f(x) \cdot L_F^m(y) + L_F^m(x) \cdot L_M^m(y). \qquad (2)
\end{aligned}$$

It should be noted that the addition theorems are not manifestly symmetric in x and y, which is a consequence of the fact that the selected and unselected markers are not interchangeable: they determine an ordered sequence. Nevertheless, since the recombination events are assumed to be randomly distributed, the association functions must be invariant under an inversion of the marker sequence. That is, the value of the function is unchanged if we interchange both the subscript and superscript as well as x and y. This implies the following inversion theorems:

$$\begin{aligned}
L_M^m(x, y) &= L_M^m(y, x) \\
L_M^f(x, y) &= L_F^m(y, x) \\
L_F^f(x, y) &= L_F^f(y, x) \\
L_F^m(x, y) &= L_M^f(y, x). \qquad (3)
\end{aligned}$$

By substitution of the addition theorems into equations 3, we obtain the following

useful relations:

$$L_M^m(x)\cdot L_M^f(y) + L_M^f(x)\cdot L_F^f(y) = L_F^f(y)\cdot L_F^m(x) + L_F^m(y)\cdot L_M^m(x)$$
$$L_F^f(x)\cdot L_F^m(y) + L_F^m(x)\cdot L_M^m(y) = L_M^m(y)\cdot L_M^f(x) + L_M^f(y)\cdot L_F^f(x)$$
$$L_M^f(x)\cdot L_F^m(y) = L_M^f(y)\cdot L_F^m(x). \qquad (4)$$

It remains to treat the case of a multiplicity of selected markers. This is particularly relevant to bacterial recombination since the recombinant chromosome is presumably circular. One would expect, then, that the origin and terminus of the donor chromosomal fragment would be discriminated against with an efficiency approaching unity, except, perhaps, when the entire chromosome is transferred. As an example, we will consider the class $\bar{a}b\bar{c}$, where $\bar{a}$ and $\bar{c}$ are selected alleles and b is unselected. If we start the Markov chain at a, then the fact that $\bar{c}$ is selected must be taken into account by normalizing to the entire class, $\overline{ac}$. In so doing, however, the chain must be begun at the same place and generated in the same direction for both classes. Therefore we have:

$$G(b \mid \bar{a}, \bar{c}) = \frac{L_F^m(x_b - x_a)\cdot L_M^f(x_c - x_b)}{L_F^f(x_c - x_a)}. \qquad (5)$$

The generalization to an arbitrary number of selected and unselected markers is clear. We could have chosen to begin the chain at c and to proceed in the opposite direction. The resulting expression can be shown to be equal to that of equation 5 by use of the third of equations 4.

DERIVATION OF THE ASSOCIATION FUNCTIONS

Absence of Lethal Lesions

By specifying the form of the association functions for sufficiently small values of their arguments, we can obtain the functions themselves from the addition theorems. For the sake of simplicity we shall postpone the inclusion of lethal effects until the following section.

A recombinant chromosome can be considered to be a linear sequence of male and female genetic information. Since this oriented sequence is assumed to be the result of a Markov process, the probability per unit distance of a transition from one source of information to the other is a constant which depends only upon the initial source. Consequently, in the limit of small x we have:

$$\left.\begin{aligned} L_M^m(x) &\approx 1 - \nu_{mr}x \\ L_M^f(x) &\approx \nu_{mr}x \end{aligned}\right\} \nu_{mr}x \ll 1$$
$$\left.\begin{aligned} L_F^f(x) &\approx 1 - \nu_{fr}x \\ L_F^m(x) &\approx \nu_{fr}x \end{aligned}\right\} \nu_{fr}x \ll 1 \qquad (6)$$

where ν_{mr} and ν_{fr} are recombination constants. These equations correspond to the familiar result that the recombination frequency between two closely spaced markers is proportional to their separations (in the absence of interference, i.e. nonstochastic effects). If we consider the first order changes in the association functions as x goes to $x + \Delta x$, we can obtain the differential equations for the functions by letting y equal Δx in equations 2, substituting equation 6 for small values of the arguments, and taking the appropriate limits as Δx goes to zero. The resulting equations are:

$$\begin{aligned} dL_M{}^m/dx &= -\nu_{mr}L_M{}^m + \nu_{fr}L_M{}^f \\ dL_M{}^f/dx &= \nu_{mr}L_M{}^m - \nu_{fr}L_M{}^f \\ dL_F{}^f/dx &= -\nu_{fr}L_F{}^f + \nu_{mr}L_F{}^m \\ dL_F{}^m/dx &= \nu_{fr}L_F{}^f - \nu_{mr}L_F{}^m. \end{aligned} \tag{7}$$

Subject to the boundary conditions implied by equations 6, as x goes to zero, these equations have the solution:

$$\begin{aligned} L_M{}^m &= \frac{\nu_{fr}}{\nu_{fr} + \nu_{mr}}\left[1 + \frac{\nu_{mr}}{\nu_{fr}}e^{-(\nu_{fr}+\nu_{mr})x}\right] \\ L_M{}^f &= \frac{\nu_{mr}}{\nu_{fr} + \nu_{mr}}[1 - e^{-(\nu_{fr}+\nu_{mr})x}] \\ L_F{}^f &= \frac{\nu_{mr}}{\nu_{fr} + \nu_{mr}}\left[1 + \frac{\nu_{fr}}{\nu_{mr}}e^{-(\nu_{fr}+\nu_{mr})x}\right] \\ L_F{}^m &= \frac{\nu_{fr}}{\nu_{fr} + \nu_{mr}}[1 - e^{-(\nu_{fr}+\nu_{mr})x}]. \end{aligned} \tag{8}$$

These equations have been obtained previously by T. H. Wood (unpublished) from an explicit analysis of the copy-choice model of recombination.

For values of x sufficiently large that the exponentials may be neglected, the association functions become:

$$\left.\begin{aligned} L_M{}^m(\infty) &= L_F{}^m(\infty) = \frac{\nu_{fr}}{\nu_{fr} + \nu_{mr}} \\ L_M{}^f(\infty) &= L_F{}^f(\infty) = \frac{\nu_{mr}}{\nu_{fr} + \nu_{mr}} \end{aligned}\right\} (\nu_{fr} + \nu_{mr})x \gg 1.$$

These asymptotic values lie between zero and unity, but in general need not be one-half. Only in the symmetric case, $\nu_{fr} = \nu_{mr}$, are they equal to one-half.

It should be noted that at sufficiently large distances, i.e. $(\nu_{fr} + \nu_{mr})x \gg 1$, the as-

sociation function values depend only upon the allele of the *unselected* marker, and not upon that of the selected marker. This loss of memory, characteristic of a stochastic process, corresponds to what is properly meant by saying that two markers are unlinked. This brings us to the following general principle: *The test of linkage, or lack of linkage, between two genetic markers is the equality, or inequality, of the association frequencies between one allele of the unselected marker and each allele of the selected marker.* In particular, no conclusions about the linkage of two genes may be drawn from the measured value of one association frequency unless other assumptions are made (e.g., symmetry). For example, to demonstrate that markers a and b are linked, it is necessary to measure the fraction of all $\bar{b}$ recombinants that are a as well as the fraction of all b recombinants that are a. If these fractions are unequal the markers are linked, regardless of the values of the fractions themselves.

The actual value of the recombination frequency depends on which pair of alleles of the selected and unselected markers is chosen. Only in the symmetric case are the two recombination frequencies equal. In fact, the measurement of these two recombination frequencies is a convenient and sensitive test for symmetry between the two parental sources of genetic information when the markers at one's disposal are closely linked.

Incorporation of Lethal Lesions

A recombinant chromosome is normally not observed directly, but rather indirectly through the phenotype of the living organism, or progeny thereof, whose genotype it determines. Lethal configurations of the DNA may arise from accidents, so to speak, occurring during the recombination process itself, or from potentially lethal mutagenic agents acting on the parental DNA prior to recombination (Wood and Marcovich, 1964). Since in genetic studies one usually deals with viable organisms, the association functions should reflect the additional contingent probability that the recombinant chromosome should contain no lethal lesions.

This may be accomplished within a stochastic framework by assigning mean densities of lethal lesions to the female and male regions of the recombinant chromosome, denoted by ν_{fl} and ν_{ml}, respectively. We may proceed to derive the differential equations for the association functions as we did earlier, but the limiting forms of the linkage functions for small values of their arguments will be modified. The reason that the recombination functions are not affected to first order in their arguments is that the probability of a recombination event is itself proportional to the distance in lowest order. For the linkage functions, we multiply the probability that no recombination event should occur by the probability that no lethal event should occur. That is:

$$L_M{}^m \approx (1 - \nu_{mr}x)(1 - \nu_{ml}x)$$
$$L_F{}^f \approx (1 - \nu_{fr}x)(1 - \nu_{fl}x).$$

To first order in x we obtain:

$$\left.\begin{aligned} L_M{}^m &\approx 1 - (\nu_{mr} + \nu_{ml})x \\ L_M{}^f &\approx \nu_{mr}x \\ L_F{}^f &\approx 1 - (\nu_{fr} + \nu_{fl})x \\ L_F{}^m &\approx \nu_{fr}x \end{aligned}\right\} \nu_{mr}x,\ \nu_{ml}x,\ \nu_{fr}x,\ \nu_{fl}x \ll 1. \tag{9}$$

From these equations it can be seen that, for closely linked markers, lethal lesions do not measurably affect the recombination frequencies $L_M{}^f$ and $L_F{}^m$, but do decrease the apparent linkage frequencies $L_M{}^m$ and $L_F{}^f$ in direct proportion to their average density. This effect provides an experimental method for measuring separately the effects of external treatments (e.g. mutagens) on the number of recombination events and the incorporation of lethal lesions into the recombinant chromosome.

Once again we let y equal Δx, substitute the limiting forms above into equations 2, and take the appropriate limits to obtain:

$$\begin{aligned} dL_M{}^m/dx &= -(\nu_{mr} + \nu_{ml})L_M{}^m + \nu_{fr}L_M{}^f \\ dL_M{}^f/dx &= \nu_{mr}L_M{}^m - (\nu_{fr} + \nu_{fl})L_M{}^f \\ dL_F{}^f/dx &= -(\nu_{fr} + \nu_{fl})L_F{}^f + \nu_{mr}L_F{}^m \\ dL_F{}^m/dx &= \nu_{fr}L_F{}^f - (\nu_{mr} + \nu_{ml})L_F{}^m \end{aligned} \tag{10}$$

The secular determinant of these equations yields a doubly degenerate pair of roots given by

$$\alpha = b/2 - [(b/2)^2 - c]^{1/2}$$

$$\beta = b/2 + [(b/2)^2 - c]^{1/2}$$

where:

$$b \equiv \nu_{fr} + \nu_{mr} + \nu_{fl} + \nu_{ml}$$

$$c \equiv \nu_{fr}\nu_{ml} + \nu_{mr}\nu_{fl} + \nu_{fl}\nu_{ml}\,.$$

The solutions we seek are those linear combinations of the general solutions of equation 10, $e^{-\alpha x}$ and $e^{-\beta x}$, which satisfy the boundary conditions implied by equations 9:

$$\begin{aligned} L_M{}^m(0) &= 1; & (dL_M{}^m/dx)_0 &= -(\nu_{mr} + \nu_{ml}) \\ L_M{}^f(0) &= 0; & (dL_M{}^f/dx)_0 &= \nu_{mr} \\ L_F{}^f(0) &= 1; & (dL_F{}^f/dx)_0 &= -(\nu_{fr} + \nu_{fl}) \\ L_F{}^m(0) &= 0; & (dL_F{}^m/dx)_0 &= \nu_{fr}\,. \end{aligned}$$

The appropriate solutions, valid for all $x \geqq 0$, are given by:

$$
\begin{aligned}
L_M{}^m &= \tfrac{1}{2}(1 + \delta)e^{-\alpha x} + \tfrac{1}{2}(1 - \delta)e^{-\beta x} \\
L_M{}^f &= \tfrac{1}{2}\nu_{mr}\gamma(e^{-\alpha x} - e^{-\beta x}) \\
L_F{}^f &= \tfrac{1}{2}(1 - \delta)e^{-\alpha x} + \tfrac{1}{2}(1 + \delta)e^{-\beta x} \\
L_F{}^m &= \tfrac{1}{2}\nu_{fr}\gamma(e^{-\alpha x} - e^{-\beta x})
\end{aligned} \tag{11}
$$

where:

$$
\begin{aligned}
\delta &\equiv \frac{\nu_{fr} - \nu_{mr} + \nu_{fl} - \nu_{ml}}{2[(b/2)^2 - c]^{1/2}} \\
\gamma &\equiv [(b/2)^2 - c]^{-1/2}.
\end{aligned}
$$

By allowing the densities ν_{fl} and ν_{ml} to go to zero, one may show that these expressions reduce to those of equations 8. Since β is always greater than α, the terms involving $e^{-\beta x}$ may be neglected for sufficiently large values of x. In this case each association function goes to zero with increasing x as $e^{-\alpha x}$, while their ratios remain constant.

From equation 11 we note that for any pair of markers the ratio of recombination frequencies for each pair of alleles, $L_M{}^f/L_F{}^m$, is just the ratio ν_{mr}/ν_{fr}. This is true regardless of their separation or the presence of any lethal effects and therefore provides a general test for symmetry.

The character of equations 11 varies greatly as the ν's take on different relative values, but there is one important set of experimental conditions for which we shall reduce them to a more simple form. This is the familiar symmetric case, $\nu_{fr} = \nu_{mr} \equiv \nu_r$, with exposure prior to mating of one of the parents, say the male, to a potentially lethal mutagenic agent. Under normal experimental circumstances the density of lethal lesions is considerably less than the density of normal recombination events, so that we may drop terms of order $(\nu_{ml}/\nu_r)^2$ compared to unity. With this approximation, and setting ν_{fl} equal to zero, equation 11 becomes

$$
\begin{aligned}
L_M{}^m &= \frac{1}{2}\left(1 - \frac{\nu_{ml}/2}{\nu_r}\right)e^{-\nu_{ml}x/2} + \frac{1}{2}\left(1 + \frac{\nu_{ml}/2}{\nu_r}\right)e^{-2\nu_r x}\, e^{-\nu_{ml}x/2} \\
L_M{}^f &= L_F{}^m = \tfrac{1}{2}(1 - e^{-2\nu_r x})\, e^{-\nu_{ml}x/2} \\
L_F{}^f &= \frac{1}{2}\left(1 + \frac{\nu_{ml}/2}{\nu_r}\right)e^{-\nu_{ml}x/2} + \frac{1}{2}\left(1 - \frac{\nu_{ml}/2}{\nu_r}\right)e^{-2\nu_r x}\, e^{-\nu_{ml}x/2}.
\end{aligned}
$$

The frequencies of surviving recombinants, $L_M{}^f$ and $L_F{}^m$, decrease with lethal density and marker separation in a simple way: the classical expression, $\tfrac{1}{2}(1 - e^{-2\nu_r x})$,

is just multiplied by the probability, $e^{-\nu_m l x/2}$, that no lethal lesions should occur in the average amount of male information incorporated. However, there is no such simplicity in the expressions for the apparent marker linkage frequencies, $L_M{}^m$ and $L_F{}^f$, where an essential asymmetry is introduced by the lethal lesions.

DISCUSSION

After calculating the association functions from the physical specifications of a particular model of recombination, one might hope to establish its validity by comparing its predictions to the observed frequencies of occurrence of genotype classes among recombinants. This, however, is not possible. Between two genes there are exactly four association frequencies, a priori. The discussion surrounding the derivation of equation 11 has stressed the biological significance of the parameters involved. However, these equations have a more general significance. From a mathematical viewpoint, we are dealing with a Markov process among four states: viable male, moribund male, viable female, and moribund female. The association functions are uniquely determined by specifying their boundary values together with the transition probabilities between the states. These probabilities are specified by four constants. The constants in equation 11 are a particular decomposition of these transition probabilities, but the forms of the equation are the most general ones possible. Consequently, any model of recombination which assumes randomness in the spatial distribution of recombination, lethal, or any other events will yield these same equations with a particular definition or rearrangement of the constants. It follows that the predictions for genetic recombination of any stochastic model are entirely independent of its physical specifications, as can be seen in the three foregoing examples:

(*a*) T. H. Wood (unpublished) has shown that a detailed analysis of the copy-choice model leads directly to equation 8.

(*b*) Equation 1, derived by Haldane (1919) from the simple break-reunion model, is equivalent to equation 8 for the symmetric case ($\nu_{mr} = \nu_{fr}$).

(*c*) The equations derived by Verhoef and de Haan (1966) for their model are identical with those of the copy-choice model and to the first and second of equation 8, apart from differences in notation.

Comparison of model predictions with experimental data permits an evaluation of the distribution moduli of sets of random events and the goodness-of-fit. If the goodness-of-fit is satisfactory, it is then possible to make negative statements about whole classes of physical models. The results of Verhoef and de Haan (1966) provide a good example of this. Since the goodness-of-fit of the association functions with their experimental data was quite satisfactory, their results (*a*) suggest that the nature of the recombination process is indeed stochastic at a macroscopic level and (*b*) deny the validity of the whole class of intrinsically symmetric models (e.g. the simple break-reunion model) for bacterial recombination.

The author is indebted to T. H. Wood, whose experiments and theoretical analysis of the copy-choice model were the impetus for this work.

This work was supported by the Institutes of National Health under Grant 5 ROI RH00501-15.

Received for publication 14 October 1968.

REFERENCES

BAILEY, N. T. J. 1961. Mathematical Theory of Genetic Linkage. The Oxford University Press, London, England.

HALDANE, J. B. S. 1919. *J. Genet.* **8**:299.

JACOB, F., and E. L. WOLLMAN. 1961. Sexuality and the Genetics of Bacteria. Academic Press, Inc., New York.

VERHOEF, C., and P. G. DE HAAN. 1966. *Mutat. Res.* **3**:101.

WOOD, T. H., and H. MARCOVICH. 1964. *Genetics.* **49**:779.

WU, T. T. 1967. *J. Theor. Biol.* **17**:40.