THE MECHANISM OF GENETIC RECOMBINATION IN PHAGE[1,2]

N. VISCONTI AND M. DELBRÜCK

Department of Genetics, Carnegie Institution of Washington, Cold Spring Harbor, Long Island, and California Institute of Technology, Pasadena

Received March 7, 1952[3]

WHEN a bacterium is mixedly infected with several compatible and genetically marked phage particles the progeny generally contains particles which show a combination of markers derived from different parental particles. The proportion of recombinant particles in the progeny is often as high as 40%. In a series of thorough studies of phage T2, HERSHEY and ROTMAN (1948, 1949) established certain basic features concerning this recombination phenomenon. These studies culminated in the recognition that phage genetics conforms in large measure to the principles worked out on higher organisms, giving evidence of linkage groups and permitting the construction of genetic maps. Recently, DOERMANN and HILL (1953b) have carried through a similar study for phage T4, and their findings fully corroborate the basic ideas developed by HERSHEY and ROTMAN. However, HERSHEY and ROTMAN also showed that a mixed infection of a bacterium is *not* a straightforward analogue of a simple genetic cross, since the progeny may contain particles which contain genetic markers derived from more than two parental particles. The interpretation of the findings, therefore, requires some generalization of the idea of a simple cross. Up till recently, the search for this generalization was strongly influenced by the idea that any mechanism proposed to explain recombination should also be adequate to explain the phenomenon of multiplicity reactivation (LURIA 1947; LURIA and DULBECCO 1949), i.e., the phenomenon that bacteria infected with several particles inactivated by ultra-violet light produce viable phage offspring in a large proportion of cases. It was thought that we are here dealing with the elimination by genetic recombination of lethal genetic factors produced by the irradiation. It is now known that multiplicity reactivation cannot be explained on this basis, and indeed, may have nothing to do with genetic mechanisms (DULBECCO 1952a).

To incorporate multiplicity reactivation into the genetic picture, one had been forced to assume that the parental particles break up into subunits which multiply independently of each other and are later reassembled into complete particles. If it is *not* required to account for multiplicity reactivation by a

[1] The experiments reported in this paper were initiated and executed by N. Visconti. Joint discussions of the results of these experiments prompted the development of the theory by M. Delbrück.

[2] Aided by grants from the National Foundation for Infantile Paralysis and from the National Institutes of Health, U. S. Public Health Service.

[3] Published out of order due to delays in editorial processing.

genetic mechanism, one may attempt to explain the genetic findings by a less radical generalization of the idea of the genetic cross, namely by the idea that mating occurs *repeatedly* during every intrabacterial cycle of phage growth. The interpretation of mixed infections thus becomes a problem in population genetics. In the present paper we propose to carry through such an interpretation, both of experiments previously published, and of some new experiments specifically designed to test the underlying assumptions.

To be able to formulate the problem properly, we must first look at certain features of the life cycle of a phage particle. It can be demonstrated by several lines of evidence that upon entering the bacterium, the phage particle undergoes a profound modification into another state in which it is unable to infect other bacteria. We will call this stage the vegetative phase.

When infected bacteria are broken up at various times during the latent period between infection and lysis, no mature particles are found during the first half of the latent period (the eclipse period). Thereafter, mature phage particles increase in number, for a few minutes slowly, and then almost linearly, up to the time of lysis (DOERMANN 1948). Several lines of evidence indicate that multiplication of vegetative phage has progressed a great deal before the appearance of the first mature bacteriophage particle. It is unknown which particular intrabacterial factors are responsible for the onset of maturation. We will make the specific assumptions, that (a) vegetative particles multiply up to the normal time of lysis, and, if lysis is inhibited, even beyond this time, (b) mature phage particles within their mother cell do not revert to vegetative phage, do not multiply, and do not genetically mix, (c) vegetative particles mate pair-wise and at random with respect to partners. With regard to the rate of mating, we assume that it increases rapidly with the concentration of vegetative particles, so that the rate becomes appreciable only after considerable multiplication has taken place. (With respect to orderliness or randomness of mating with respect to time, we will introduce specific assumptions later on.) By way of justifying these assumptions, we cite the following facts:

(1) Two lines of evidence suggest that it is necessary to assume several rounds of mating. One of these is the occurrence of recombinants combining genetic markers from three different parents. A second line of evidence is provided by the finding (DOERMANN and HILL 1953a) that in a two-factor cross with unequal multiplicity of infection by the two parental particles the progeny contains more recombinants than particles with the genotype of the minority parent. This indicates that the genetic mixing has gone farther toward genetic equilibrium than could be accomplished by a single round of mating.

(2) Two lines of evidence indicate that there is a drift *in the course of time* towards genetic equilibrium. On the one hand, in crosses involving *unlinked markers* it is found (DOERMANN and HILL 1953a) that the fraction of recombinants in the yield increases from 34% to 42% when the mature progeny is sampled by artificial premature lysis at various times during the second half of the latent period. In two cases of linked factors the fraction increases from

6.0% to 11.5% and from 25% to 37%. The genetic equilibrium lies at 50% recombinants, and the approach to this equilibrium during the period investigated by DOERMANN can be satisfactorily accounted for by our theory, as will be shown elsewhere. On the other hand, in crosses involving *closely linked markers,* the frequency of recombinants can be greatly increased by inhibiting lysis for a period several times the length of the latent period. During this prolonged intracellular phase, the proportion of recombinants among the mature particles increases approximately linearly (LEVINTHAL and VISCONTI 1953).

(3) The assumption that most of the recombination takes place after a great deal of multiplication has taken place is based on the finding (HERSHEY and ROTMAN 1949; DOERMANN and HILL 1953a), that in crosses involving closely linked factors the recombinants are approximately randomly distributed among the yields obtained from individual bacteria. If recombination occurred *exclusively* very early, before appreciable multiplication had taken place, one should have found, instead, large clones of recombinants from any bacteria in which recombination occurred at all.

(4) When the progeny of phage from individual bacteria is examined (HERSHEY and ROTMAN 1949) little or no correlation between the numbers of opposite type recombinants is found. We interpret this lack of correlation not as indicating that the two opposite types are produced in statistically independent elementary acts, but as a reflection of secondary events (random multiplication and random maturation after recombination) which obscure an original correlation which may be perfect in every mating.

(5) The most important assumption of our theory is the assumption that genetic mixing occurs as a result of a series of elementary acts (matings) in each of which a pair of *genetically complete* phage particles is involved. An interesting implication of this assumption is a statistical one: it leads to positive correlations between different exchanges affecting the same particles of the progeny, even if these exchanges have zero correlation for each individual mating. Qualitatively this may be seen by an extreme example: suppose the progeny consisted of a mixture of two populations, one of which had not mated at all, while the other had mated numerous times. It is clear that in the mixture of these two partial populations different exchanges would be positively correlated although there may be no correlation within each partial population. In our theory the final population can be looked upon as a mixture of populations with different degrees of mating experience, leading to such positive correlations between different exchanges. This type of correlation had been apparent in some of the two- and three-factor crosses of HERSHEY and ROTMAN (1948, 1949). Several experiments in the present paper specifically designed to bring out this point show the effect more clearly. If one were to interpret recombination values as in ordinary genetics, with disregard of the complications resulting from repeated mating, one would call this positive correlation a " negative interference," and would have good reason to wonder why one " cross-over " should *increase* the probability of a second one in the same linkage group.

Our problem may now be stated in the following terms: the parental particles enter the cell and immediately become vegetative particles. These multiply, and after they have multiplied up to around, say 40, the rate of mating becomes appreciable. Soon after this, particles begin to be withdrawn from this pool of mating and multiplying particles, and are set aside as mature particles. The progeny examined at the time of lysis, therefore, consists of a mixture of samples withdrawn at various times from the mating pool. This withdrawal occurs at an approximately linear rate. In comparing our theory with experimental results we will for the most part make the idealizing assumption that the progeny examined is one that has been withdrawn in its entirety *at a particular moment* from the mating pool. The real progeny will differ from this idealized one in a manner which will be discussed after presenting the theory.

THEORY

Most of the new facts to be reported in this paper deal with crosses in which there are two parental types which differ with respect to three genetic factors, either unlinked, or linked in various ways. Another variable that can be controlled is the multiplicity of infection. Here the two cases of interest are, (a) equal, high multiplicity of infection of both parents, and (b) high multiplicity of infection with one parent, and single infection with the other parent.

The chief advantage of working with three-factor crosses as compared to two-factor crosses is this: in three-factor crosses we can focus our attention on those particles in the yield which are recombinants with respect to two of the factors, and then ask in what fraction of these there has also occurred a recombination with respect to the third factor. In this way we can be certain that our genetic ratios are not falsified by the inclusion of unadsorbed parental particles or by particles derived from bacteria that did not get mixedly infected, or by particles which for some other reason failed to meet a parental particle of the opposite type. On the other hand, the calculation of the influence of matings on intra-bacterial phage populations involving three factors (eight genotypes) is somewhat involved, and requires algebraic analysis. This we will now proceed to develop.

Synchronous mating and random-in-time mating

We calculate first the effects of n successive, nonoverlapping rounds of random mating on an arbitrary mixture of genotypes, assuming that every phage particle participates in every round of mating. This we will call the case of *synchronous* mating. This assumption may be too crude, and we would like to see how the conclusions have to be modified if mating does not occur in successive, nonoverlapping rounds, but for each particle with a constant probability per time unit. In this case, after an interval of time in which each particle has mated on the average m times, there will be a Poisson distribution, $p_r = m^r e^{-m}/r!$, of the number of particles that have mated during this time 0, 1, . . . r . . . , times. It must be noted that, under such a regime, which we will call *random-in-time* mating, the particles which mate just once do not all

mate within a population that has the original genetic composition. Some may mate towards the end of the period when the genetic composition of the population into which they mate is entirely different from the original one, and, similarly, for the particles which mate more than once, the effects might be expected to be quite different from those of nonoverlapping rounds of mating. However, in the appendix we will prove the important theorem that in the case of two- or three-factor crosses (but not when more than three factors are involved!) the effect of this regime is the same as if the total population consisted of fractions $p_0, p_1, p_2, \ldots$, which undergo 0, 1, 2, ... successive nonoverlapping rounds of matings, *each group starting with the original genetic composition of the total population and mating within itself.*

The transition from " synchronous " to " random-in-time " mating can now be made with ease because all the genotype frequencies for a fixed number n of rounds of matings turn out to be linear combinations of the form $A + Bb^n$ (see equation 8) and the average of such an expression for a Poisson distribution of the number n is given by

$$\Sigma_n p_n (A + Bb^n) = A + Be^{-m(1-b)}. \qquad (1)$$

We need not make any specific assumptions as to whether or not multiplication takes place during recombination. This is so because all calculations refer to the *proportions* of the various genotypes, and these proportions are not altered by reproduction.

Our treatment will follow closely the methods developed by GEIRINGER (1944) in connection with a more general problem, involving an arbitrary number of factors, an arbitrary number of alleles for each factor, and arbitrary linkages. In the three-factor case her method can be presented in a somewhat simpler and more visualizable form.

Three-factor crosses

We designate a given genotype by the symbol (ijk), where the three letters represent the three genetic factors, and each of them may assume the value 1 or 2 for the two alleles of the factor.

In figure 1 the eight genotypes are represented by the eight corners of a cube as explained in the legend.

The genotypes differing from a given one in a *single* factor are its neighbors along the three edges which meet at the corner representing this type. Types at opposite ends of a face diagonal differ in two factors, those at opposite ends of a space diagonal in all three factors. An allele frequency is obtained by taking a sum around a face of the cube, and the frequency of a certain combination of two alleles, by taking the sum along an edge.

In the cross used by HERSHEY and ROTMAN to demonstrate the occurrence of triparental recombinants four corners of a regular tetrahedron are involved.

We imagine that the mating between two particles of type (ijk) and (lmn) leads to the formation of a zygote which immediately segregates again into haploid particles. This segregation can occur in four possible ways, namely,

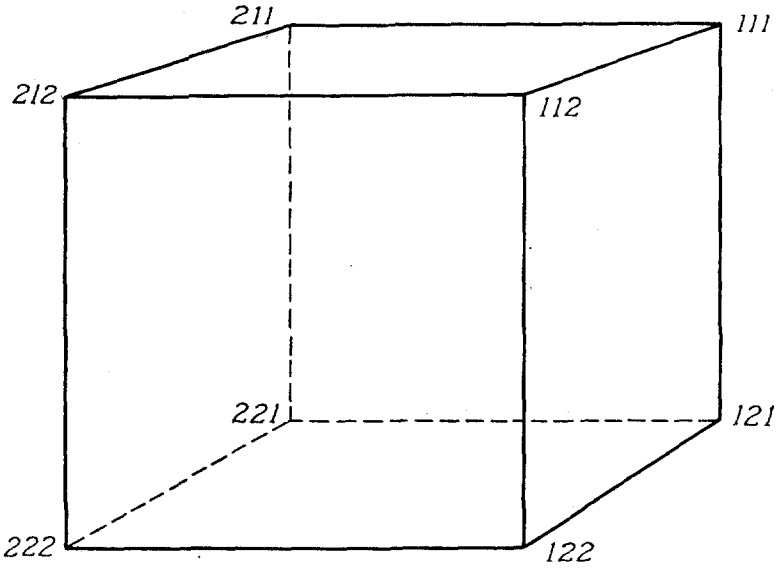

FIGURE 1.—The eight genotypes are represented by the eight corners of a cube in such a fashion that all the corners of the right face of the cube represent the genotypes carrying the "1" allele of the first factor, and the corners of the left face the "2" allele of this factor. Similarly, top and bottom face correspond respectively, to the "1", and "2" allele of the second factor, and back and front face to those of the third factor.

(0) into the parental types, (1) such that factor 1 has been segregated from 2 and 3, (2) such that factor 2 has been segregated from 1 and 3, and (3) such that factor 3 has been segregated from 1 and 2. We will designate the probabilities of these four alternative events with c_0, c_1, c_2, and c_3, the sum of which is equal to unity.

The probabilities c_r are related in a simple manner to the probabilities p_{rs} of recombination between any two factors, namely,

$$c_r + c_s = p_{rs} \tag{2}$$

Table 1 lists the values of the coefficients c_i for four types of linkage relationships in which we will be particularly interested. For the cases of linkage

TABLE 1

The recombination parameters c for particular linkage relationships.

Type of linkage	c_0	$c_0 + c_1$	$c_0 + c_2$	$c_0 + c_3$
All unlinked	1/4	1/2	1/2	1/2
(1, 2) linked, 3 unlinked	$(1 - p_{12})/2$	1/2	1/2	$1 - p_{12}$
All three linked, order 1, 2, 3 (no interference)	$1 - p_{12} - p_{23}(1 - p_{12})$	$1 - p_{23}$	$1 - p_{12} - p_{23}(1 - 2p_{12})$	$1 - p_{12}$
All three linked, order 1, 2, 3 (complete interference)	$1 - p_{12} - p_{23}$	$1 - p_{23}$	$1 - p_{12} - p_{23}$	$1 - p_{12}$

between all three factors it has been assumed that the factors lie in the order 1, 2, 3. With respect to interference between recombinations between factors 1 and 2 and between factors 2 and 3, we list the two cases of no interference and of complete interference. No interference implies the relation

$$p_{13} = p_{12}(1 - p_{23}) + p_{23}(1 - p_{12}) \tag{4}$$

while for complete interference

$$p_{13} = p_{12} + p_{23}. \tag{4a}$$

We designate the frequency of the genotype (ijk) after n rounds of mating by $a_{ijk}^{(n)}$. The same symbol without the superscript designates the original genotype frequencies, i.e., the parental mixture. By definition the sum of the frequencies is unity at any time.

Our equations will involve certain sums of genotype frequencies, and for these we will introduce convenient symbols. We designate the sum of the frequencies of those genotypes which carry a particular allele of a particular factor by an *a* in the subscripts of which the indices representing the unspecified factors have been replaced by dots. Thus $a_{..k}$ designates the frequency in the population of the k allele of the third factor. These allele frequencies are not affected by mating.

Similarly we designate by $a_{.jk}$ the sum of the frequencies of the genotypes carrying the j allele of the second factor and the k allele of the third factor. This sum is the frequency in the population of this particular allele combination, and it will, in general, change with each round of mating.

Similarly, we designate the sum of all genotype frequencies by $a_{...}^{(n)} = 1$ (for any n).

We now wish to calculate the effect of one round of mating on the frequency of a particular genotype. There will be contributions to this genotype from a variety of zygotes and we will group these contributions according to the type of segregation that gives rise to it. In order that a given type of segregation, say c_1 (that in which the first factor is segregated from the others), give rise to a given genotype, say (ijk), the zygote must contain the i allele of the first factor at least once, and similarly the j allele of the second factor, and the k allele of the third factor. Moreover, these alleles must occur in the zygote so that the segregation of type c_1 can give rise to this genotype, i.e., the zygote must be one that has been formed from two particles one of which had the constitution (i··) and the other one the constitution (·jk), where the dots represent arbitrary alleles. These and only these zygotes will give the desired genotype by means of the specified type of segregation.

We must now consider the probability that a given zygote is of a particular constitution, this probability to be expressed in terms of the relative frequencies, a_{ijk}, of the haploid genotypes. In the case of homozygotes, (ijk)×(ijk), this probability is simply a_{ijk}^2. In the case of heterozygotes, (ij'k')×(i'jk), the probability is *twice* the product of the genotype frequencies, since the sequence of the conjugants is irrelevant.

In calculating the contribution of these zygotes to the genotype (ijk) we

must next take into account that the homozygotes contribute all of their segregants to (ijk) while the heterozygotes contribute only half of their segregants.

The total contribution from both homozygotes and heterozygotes, by segregation of type c_1, to the genotype (ijk) is thus simply

$$c_1 \Sigma_{i',j',k'} (a_{ij'k'} a_{i'jk}) \tag{5}$$

Since none of the indices over which the sum is to be extended occurs in both factors we may perform the sum for each factor separately and obtain the very simple expression $c_1 a_i .. a_{.jk}$.

The effect of one round of random mating on the frequencies of the genotypes can thus be expressed by the following equation:

$$a_{ijk}^{(n)} = c_0 a_{ijk}^{(n-1)} a_{...}^{(n-1)} + c_1 a_{i..}^{(n-1)} a_{.jk}^{(n-1)} + c_2 a_{.j.}^{(n-1)} a_{i.k}^{(n-1)} + c_3 a_{..k}^{(n-1)} a_{ij.}^{(n-1)} \tag{6a}$$

This recurrence relation expressing the genotype frequencies must now be "integrated," i.e., $a_{ijk}^{(n)}$ is to be expressed in terms of the original genotype frequencies. To do this we observe first that $a... = 1$ and that the symbols involving two dots represent allele frequencies and are therefore constants. The symbols involving a single dot are not constants, but they obey a simpler recurrence relation which can be obtained from equation (6a) by summing over one of the indices, thus, for instance, putting $n = 1$,

$$a_{.jk}^{(1)} = c_0 a_{.jk} + c_1 a_{...} a_{.jk} + c_2 a_{.j.} a_{..k} + c_3 a_{..k} a_{.j.} = (c_0 + c_1) a_{.jk} + (c_2 + c_3) a_{.j.} a_{..k} \tag{6b}$$

This is the general recurrence relation for a two-factor cross. We note that the second term on the right hand side of equation (6b) is a constant, since $a_{.j.}$ and $a_{..k}$ represent allele frequencies. Equation (6b) is therefore of the form

$$\begin{aligned} x_n &= Ax_{n-1} + B \\ x_{n-1} &= Ax_{n-2} + B \\ &\cdots \\ x_1 &= Ax_0 + B \end{aligned}$$

where $A = c_0 + c_1$, and $B = (c_2 + c_3) a_{.j.} a_{..k}$. This is easily "integrated" by multiplying the second equation with A, the third with A^2, etc., and adding all equations. Thus we obtain x_n in terms of x_0 and the constants:

$$x_n = A^n x_0 + B(1 - A^n)/(1 - A)$$

or, in our case,

$$a_{.jk}^{(n)} = (c_0 + c_1)^n a_{.jk} + [1 - (c_0 + c_1)^n] a_{.j.} a_{..k} \tag{7}$$

We will return to this general solution of the two-factor case later. Here we use it to substitute the general expressions for the symbols involving a single dot from equation (7) into equation (6a) which thereby assume the form

$$\begin{aligned} x_n &= Ax_{n-1} + B + CD^{n-1} \\ x_{n-1} &= Ax_{n-2} + B + CD^{n-2} \\ &\cdots \\ &\cdots \\ x_1 &= Ax_0 + B + C \end{aligned}$$

This we integrate again by multiplying the second equation by A, the third by A^2, and so on, and adding all equations:

$$x_n = x_0A^n + B(A^n - 1)/(A - 1) + CD^{n-1}[(A/D)^n - 1]/[(A/D) - 1]$$

Applying this procedure to equation (6a) we obtain

$$a_{ijk}^{(n)} = c_0^n a_{ijk} + \{a_{i..}a_{.jk}[(c_0 + c_1)^n - c_0^n] + [\ldots] + [\ldots]\} + a_{i..}a_{.j.}a_{..k}\{1 - c_0^n - \Sigma_{r=1}^{3}[(c_0 + c_r)^n - c_0^n]\} \quad (8)$$

The products involved in equation (6a) bear the following relation to the corner representing the genotype (ijk): at this corner three edges and three faces of the cube meet; the quantity multiplied into c_1 is the product of the edge along the 1 axis by the face at right angles to this axis, and similarly for the other terms. Equation (8) contains the same combinations, and in addition the last term, which is simply the product of the three faces meeting at the corner (ijk).

If infection is made with a mixture of only two genotypes, differing in all three factors, we may express the genotype frequencies through a parameter f:

$$a_{111} = (1 + f)/2,\ a_{222} = (1 - f)/2,\ \text{all others} = 0 \quad (9)$$

We will call the parents majority and minority parent, respectively. Among the recombinants we will distinguish those which differ from the majority parent by only one factor from those which differ from the minority parent by only one factor. In the case of unlinked factors the first group will be more frequent than the second group, and we will call them, correspondingly, majority and minority recombinants.

Tables 2 and 3 list the coefficients which occur in equation (8) for the special cases of the bi-parental cross described by equation (9) and the tri-parental cross mentioned in the introduction. These coefficients are easily obtained by the rules stated in the preceding paragraphs.

Equation (8) states the solution of our problem in the most general terms. We have just indicated how this is to be handled for the special cases of bi- and tri-parental crosses. Other specializations might concern the types of link-

TABLE 2

The coefficients occurring in equation (8) for the two-parental cross described by equation (9).

Type wanted	ijk	$4a_{i..}a_{.jk}$	$4a_{.j.}a_{i.k}$	$4a_{..k}a_{ij.}$	$8a_{i..}a_{.j.}a_{..k}$
Majority parent	111	$(1 + f)^2$	$(1 + f)^2$	$(1 + f)^2$	$(1 + f)^3$
Majority recombin.	211	$1 - f^2$	0	0	$(1 + f)^2(1 - f)$
	121	0	$1 - f^2$	0	
	112	0	0	$1 - f^2$	
Minority recombin.	122	$1 - f^2$	0	0	$(1 - f)^2(1 + f)$
	212	0	$1 - f^2$	0	
	221	0	0	$1 - f^2$	
Minority parent	222	$(1 - f)^2$	$(1 - f)^2$	$(1 - f)^2$	$(1 - f)^3$

TABLE 3

The coefficients occurring in equation (8) for the parental mixture $a_{211} = a_{121} = a_{112} = 1/3$, *all others* $= 0$.

ijk	$27a_{i\cdot\cdot}a_{\cdot jk}$	$27a_{\cdot j\cdot}a_{i\cdot k}$	$27a_{\cdot\cdot k}a_{ij\cdot}$	$27a_{i\cdot\cdot}a_{\cdot j\cdot}a_{\cdot\cdot k}$
111	6	6	6	8
211	3	6	6	4
121	6	3	6	4
112	6	6	3	4
122	0	3	3	2
212	3	0	3	2
221	3	3	0	2
222	0	0	0	1

age between the three factors. Table 1 lists the values of the c_i to be chosen for four particular cases. The corresponding results for the case of random-in-time mating are obtainable in every case by applying relation (1).

Two-factor crosses

We now return to equation (7) which expresses the effect of n rounds of random mating on an arbitrary initial population, and with arbitrary linkage ($c_0 + c_1 = 1 - p$) between two factors. It has been derived previously by several authors, including GEIRINGER (1944).

If infection is made with a mixture of only two genotypes, differing in both factors, we may write

$$\begin{aligned} a_{11} &= (1+f)/2 \quad \text{(majority parent)} \\ a_{22} &= (1-f)/2 \quad \text{(minority parent)} \\ a_{12} &= a_{21} = 0 \end{aligned} \tag{9a}$$

Substituting these values for a_{jk} in equation (7) we obtain for the frequencies of parentals and recombinants after n rounds of synchronous random mating

$$\begin{aligned} \text{Maj. parent} \quad & 4a_{11}^{(n)} = 2(1+f)(1-p)^n + (1+f)^2[1-(1-p)^n] \\ \text{Min. parent} \quad & 4a_{22}^{(n)} = 2(1-f)(1-p)^n + (1-f)^2[1-(1-p)^n] \\ \text{Recombinant} \quad & 4a_{12}^{(n)} = (1-f^2)[1-(1-p)^n] \end{aligned} \tag{10a}$$

For random-in-time mating, with an average of m matings, we obtain, applying equation (1),

$$\begin{aligned} \text{Maj. parent} \quad & 4a_{11}^{(m)} = (1+f)^2 + (1-f^2)e^{-mp} \\ \text{Min. parent} \quad & 4a_{22}^{(m)} = (1-f)^2 + (1-f^2)e^{-mp} \\ \text{Recombinant} \quad & 4a_{12}^{(m)} = (1-f^2)(1-e^{-mp}) \end{aligned} \tag{10b}$$

Special cases of these relations are represented in figure 4c of the experimental part.

Corrections for spread in maturation time

We mentioned in the introduction that the phage population inspected in any experiment is actually a mixture of particles withdrawn from the mating pool

by maturation at various times, and that there are strong indications that among the vegetative phages mating continues during this period of maturation. Let us call m_0 the average number of rounds of matings at the time when the first particles mature, and m_1 the average number of rounds of matings at the time of lysis.

We estimate the effect of the spread of maturation time on the genotype frequencies by the following procedure. We assume that the final population is a linear mixture of populations coming from mating pools in which m varies from m_0 to m_1. Analytically this involves the averaging of expressions of the type e^{-ma}. This average is

$$e^{-ma} = e^{-m_0 a}(1 - e^{-(m_1 - m_0)a})/a(m_1 - m_0) \tag{11}$$

This procedure introduces two parameters, m_0 and m_1 instead of the previous single m. Of these we determine m_0 by an independent experiment. DOERMANN and HILL (1953a) have measured the recombination value for two pairs of markers, assumed to be unlinked, among the earliest matured phages. This was done by experiments involving premature lysis. According to our assumptions, this should represent the frequency of the recombinant genotype at the time when the average number of rounds of mating equals m_0, and from equation (10b) this frequency should be

$$V_2 = (1 - e^{-m_0/2})/2 \tag{10c}$$

The values found by DOERMANN and HILL for two different cases of presumably unlinked factors are $V_2 = .34$ and $.25$, from which $m_0 = 2.3$ and 1.4, respectively. It is not certain that the factor pair giving $V_2 = .25$ is really unlinked. If it is slightly linked, the larger value, $m_0 = 2.3$ would have to be used. In our calculations in the later part of this paper, we base our estimates of the correction introduced by taking into account the spread of maturation time on $m_0 = 1.4$. If the actual m_0 is greater than this, the corrections would be smaller. In plotting genotype frequencies and their ratios, we use as the abscissa either n, the number of synchronous rounds of mating, or m, the average number of rounds of mating at a fixed maturation time, or $\bar{m} = (m_0 + m_1)/2$, the average number of rounds of mating when the maturation time has a spread.

MATERIALS AND METHODS

The viral mutations used as markers in our crosses were the *h* (host range) mutant, several *r* (rapidly lysing) mutants, and the *m* (minute) mutant. All these mutants were isolated from the strain H of the bacteriophage T2, and have been extensively described by HERSHEY and ROTMAN (1948 and 1949). The *h* mutant lyses strain B/2 of *E. coli* which is resistant to the h^+ type. When plated on strain B of *E. coli*, *h* and h^+ give the same type of plaque, but if a mixture of B and B/2 is used for plating, the *h* plaques appear completely clear, while the h^+ plaques appear turbid as the phage is able to lyse only one of the two types of bacteria used as indicators. The *r* and the *m* mutant give a type of plaque which is clearly different from the normal one, as described in the papers of HERSHEY and ROTMAN.

The procedure of making a cross between two different strains of phage consists in infecting a certain number of growing bacteria with the two kinds of virus. The bacteria, infected with a mixture of the two, will lyse, yielding new phage. We refer to the infecting phage as the parent, and to the phage which is liberated upon lysis as the progeny. In this work we have used crosses either with equal or extremely unequal multiplicity. In the first type, we use the two parents in equal amount, and in order to reduce random variations in the number of particles adsorbed per bacterium, a multiplicity between 15 and 20 particles per cell was used. DULBECCO (1949) has shown that not more than about 10 phage particles can take part in the growth inside a single bacterium. For this reason there is no advantage in using a higher multiplicity. The other type of cross, with very unequal multiplicity, consists in infecting the bacteria with an average of between 10 and 15 particles of one type, and an average of one phage of the other type to four bacterial cells. With this procedure we infect each cell with a high multiplicity of one of the parents (majority parent) and just a fraction of the cells with one particle of the other parent (minority parent). Few cells will adsorb more than one particle of the minority parent. To avoid exclusion of some particles due to previous infection (DULBECCO 1952b), adsorption was made in buffer using a technique described by BENZER (1952).

The technical procedure in making a cross consists of the following: a two-hour culture of *E. coli* strain H in broth, containing 10^8 cells per ml, is centrifuged, washed in buffer, and resuspended in buffer at a concentration of 2×10^8 cells per ml. The suspension is mixed with an equal volume of a suspension of phage in buffer. The bacteria mixed with the phage are kept 10 minutes at 37°C; at the end of this period more than 80% of the phage particles are adsorbed to the bacteria. At the end of the adsorption period, phage anti-serum is added to eliminate the unadsorbed phage. After 5 minutes of anti-serum treatment, the suspension is diluted in broth to a concentration of 10^3 bacteria per ml, and a tube containing this final dilution is kept at 37°C for two hours. At the end of this period all the cells have burst, and we can plate the yield after a further dilution made in order to have approximately 100 plaques per plate.

Two different methods were used to classify the different genetic types among the progeny. In one method the phage yield is plated on double indicator (B + B/2) so that every plaque can be classified for *h*, *r*, and *m*. As it is difficult to distinguish between *rm* and r^+m, these two categories have always been pooled together. We call this the direct method. The other method, which we will call the sampling method, is applied in some three-factor crosses in which two of the markers are of the *r* or *m* type, and the third is of the *h* type. In this case, by plating on single indicator (B), we select plaques which show recombination for the first two markers. These plaques can be sampled, and by streaking the sample on B/2, they can be classified as *h* or h^+. For instance, using two *r*'s at different loci, r_1 and r_2, we can make the following cross $r_1r_2^+h^+ \times r_1^+r_2h$. The r^+ plaques can be sampled and classified as *h* or h^+ by

streaking on B/2. This method is more laborious than the direct classification on double indicator plates, but it is also more reliable.

The nutrient broth used is composed of Bacto peptone 10 g, Bacto beef extract 3 g, sodium chloride 5 g, glucose 1 g, per liter distilled water. The pH is 6.8. The buffer used for adsorption is composed of $MgSO_4$ (10^{-3} M), $CaCl_2$ (10^{-4} M), $Na_2HPO_4 \cdot 12H_2O$ (7.6 g) KH_2PO_4 (1.5 g), NaCl (4.0 g), K_2SO_4 (5.0 g), gelatin (.01 g), redistilled H_2O (1000 ml), pH = 7. Nutrient agar plates are poured with a minimum of 35 ml per 10 cm Petri dish of the following mixture: Bacto agar (10 g), Bacto tryptone (10 g), NaCl (8 g), Na-citrate (2 g), glucose (1 g), H_2O (1000 ml). Agar for layer plating has the same composition, but contains only .6% agar. An amount of .2 ml of virus suspension is added to a tube containing 2 ml of top layer agar at 45°C, and 2 drops of a 24-hour culture of the indicator bacterial strain. Plates are incubated for 18 hours at 37°C before being read.

EXPERIMENTS

The experiments we now wish to report correspond to special cases of the general theory. These cases are selected from several points of view. First, we describe a method which permits a test of the assumption of nonlinkage between two factors. Second, we select a special case where the ratio between two genotypes affords a sensitive measure of the average number of rounds of matings. Third, we select a case which discriminates sharply between the assumptions of synchronous mating and random-in-time mating. Each of these cases concerns unlinked factors. Comparison with the experimental values permits us to estimate the parameters of the theory. We then turn to the consideration of bi-parental crosses involving three linked factors. Our attention will be focused on the recombination between two particular linked factors, *A* and *B*, among that portion of the progeny in which the second factor, *B*, has recombined with the third, *C*. Our theory, with its specifications derived from the analysis of unlinked factors, makes definite predictions as to how the recombination of *A* and *B* in the selected part of the progeny should vary with the degree of linkage between *B* and *C*, and with the relative multiplicity of infection. The comparison of these predictions with the experimental results constitutes the principal test of the theory.

Unlinked factors

(1) *A test of nonlinkage.* There are a number of pairs of factors whose recombinants appear in the progeny with what appears to be a maximum frequency, and we conjecture that these factors are truly unlinked, i.e., that in individual matings between particles differing in these two factors the probability of segregation into recombinants is exactly equal to one-half. This assumption cannot be tested directly because we do not have available a uniform population produced by exactly one round of mating between unequal particles. However, the assumption can be tested in three-factor crosses by comparing certain classes of recombinants. Thus, in a three-factor cross in-

TABLE 4

Results obtained in the cross hr_1m^+ (111) × $h^+r_1^+m$ (222) with equal input of the two parents. The yield of cross 1 was examined by the direct method, the yield of cross 2 by the sampling method.

(111) × (222)

	(·11)		(··2)		(·21)	
	(111)	(211)	(1·2)	(2·2)	(121)	(221)
Cross 1	142	93	220	252	105	99
Cross 2	356		627		106	112

Ratios between frequencies of different recombinants

Cross 1 (121)/(221) = 1.06 (211)/(221) = .94 Cross 2(121)/(221) = .95

volving three unlinked factors, and made with equal input of the two parental genotypes, (111) and (222), the six recombinant genotypes should all be exactly equally frequent, irrespective of the number of rounds of matings and of their distribution in time. Table 4 shows an example of such a cross, in which three of the six recombinant genotypes could be estimated individually, and these three classes are indeed equally frequent within the limits of sampling errors. However, we have to be careful in assessing the value of this test. First we have to pay attention as to which classes, precisely, were found to be equally frequent. For instance, if the classes (211) and (221) are found equally frequent this merely means that the second factor is equally linked (or unlinked) with the first and third, since in recombinants between the first and third factor, the two alleles of the second factor are equally distributed.

If now, a third recombinant genotype is found equally frequent with the two just given, say (212), this means that also the third factor is equally linked (or unlinked) with the first and second factor. Altogether, we conclude then that all three factors have a symmetric linkage relationship with each other, and this means that they are all three at least very nearly unlinked.

One would like to estimate precisely how large a deviation from nonlinkage would still be compatible with our findings. Qualitatively one can say this: the larger the number of rounds of matings, the less conspicuous will be any effect of a slight amount of linkage. Actually, there are two opposing tendencies: on the one hand the effects of residual linkage are cumulative since they occur on every round of mating. On the other hand, every round of mating brings the genotypic mixture closer to genetic equilibrium, in a rapidly converging sequence. Thus, the deviation from equilibrium, within which the residual linkage might express itself, decreases rapidly, and differences of this deviation, due to residual linkage, become less and less measurable.

In two-factor crosses between unlinked factors the frequency of recombinants is appreciably below 50 percent. From our point of view this deficit from 50 percent is largely due to the spread in maturation time and the fact that at the beginning of maturation the population is still far from genetic equilibrium. Since we cannot predict this deficit quantitatively two-factor crosses furnish no

sharp criterion for nonlinkage. The three-factor test given here predicts a definite value for certain genotype ratios in case of nonlinkage. It could be made more sensitive by inspecting the progeny at a stage when it has progressed less towards genetic equilibrium, using premature lysis.

(2) *The average number of rounds of matings.* Three unlinked factors; unequal input. In the case of equal input all six recombinants are equally frequent and their frequency is equal to $(1 - 1/2^n)/8$. This fraction approaches 1/8 very rapidly with increasing n. A much more sensitive measure of n is obtained by making the input of the two parental types very unequal. In this case we have two classes of recombinants, namely, majority recombinants, which differ by only one factor from the majority parent, and minority recombinants, which differ by only one factor from the minority parent. The frequencies of these recombinants depend on the number of rounds of matings in the following manner:

$$\text{major. recombinant } 8a_{112}^{(n)} = (1 - 1/2^n)(1 - f^2)(1 + f - 2f/2^n) \qquad (12a)$$

$$\text{minor. recombinant } 8a_{221}^{(n)} = (1 - 1/2^n)(1 - f^2)(1 - f + 2f/2^n) \qquad (12b)$$

In figure 3 these frequencies are plotted as a function of n for an input ratio of 9:1 (f = .8). The majority recombinant increases steadily with n, while the minority recombinant at first increases and then decreases. Qualitatively this behavior of the minority recombinant may be understood as follows: in the first round of mating most of the minority parents mate with majority parents. This results in a great reduction of the fraction of minority parents and in the formation of a certain fraction of minority recombinants. In the next round of mating most particles again mate with majority parents. Since there are now much fewer minority parents, there will be a much smaller contribution from this type of zygote to the minority recombinants than in the first round. On the other hand, the minority recombinants formed during the first round will be largely eliminated by recombinations formed in the matings of these particles with the majority parent. Thus, a net loss of minority recombinants occurs during the second round.

Figure 4a shows the ratio of majority recombinants to minority recombinants. This ratio reaches a limiting value equal to the input ratio when genetic equilibrium is reached (n = infinity). However, this value is approached only very gradually. Even after five rounds of mating, the deviation from genetic equilibrium is very appreciable. The corresponding calculations for random-in-time mating are shown by dashed lines in figures 3 and 4a. Here the approach to equilibrium is somewhat slower, but not sufficiently so to permit good discrimination. The corresponding calculations, if spread in maturation is taken into account according to the procedure outlined before (equation 11), are shown by the dotted line in figure 4a.

In table 5 the results of two three-factor crosses using unequal multiplicity (input 9 to 1) are shown. The two parental types differing in three unlinked characters are the same as those in the cross of table 4. The two crosses differ with respect to which of the parents was used as majority parent. The ratio

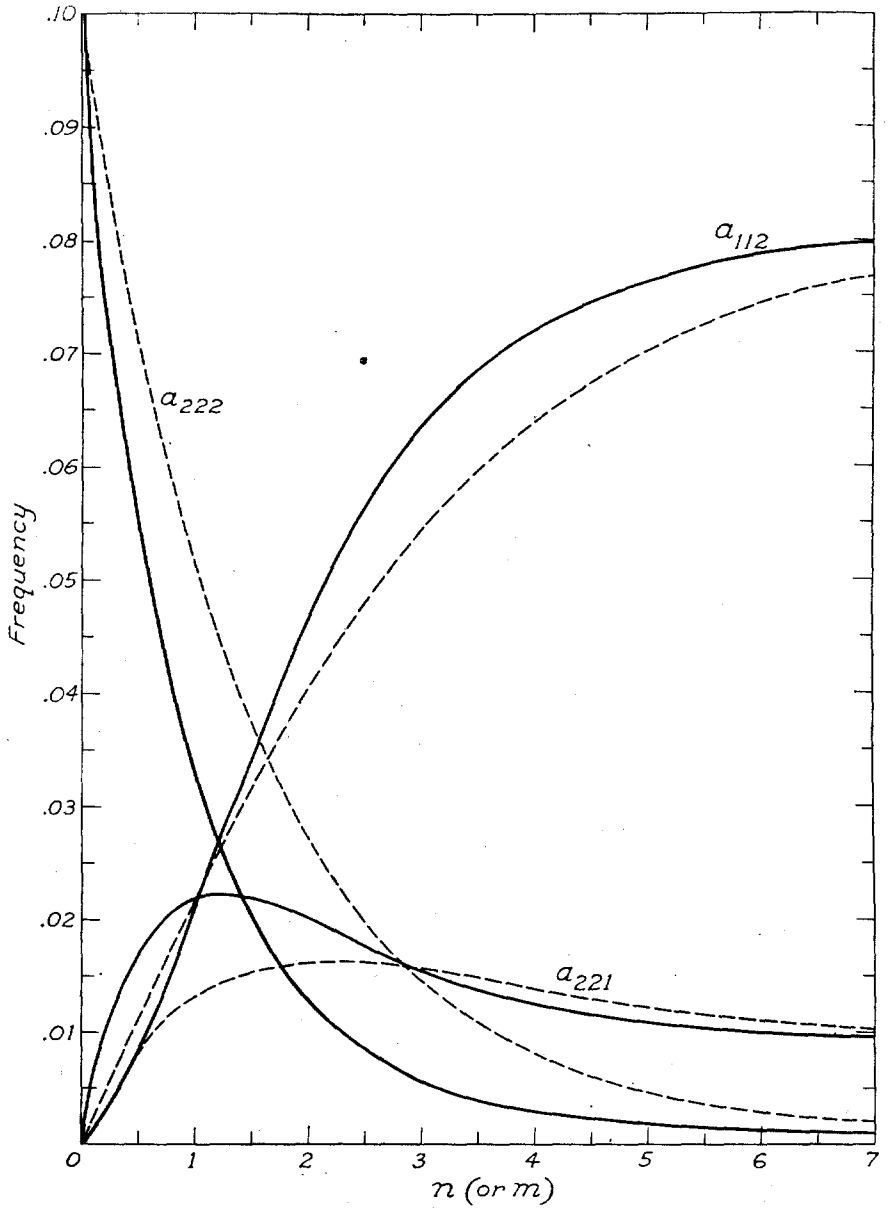

FIGURE 3.—The frequency of a majority recombinant, a minority recombinant, and the minority parent (a_{112}, a_{221}, a_{222}) in the progeny of three-factor crosses, as a function of the number of rounds of mating. The continuous curves refer to synchronous mating (n rounds), the dashed curves to random-in-time mating (m rounds). Unlinked factors.

TABLE 5

Results obtained in the cross $hr_1m^+ \times h^+r_1{}^+m$, using an excess of one parent over the other (input ratio 9 to 1). The yield was examined by the direct method. In cross 1 an excess of hr_1m^+ (111) was used, in cross 2 an excess of $h^+r_1{}^+m$ (111). In both cases we indicate the majority parent with (111).

(111) × (222)

	(111)	(121)	(112)	(122)	(211)	(221)	(212)	(222)
Cross 1	12581	288	319		289	51	86	
Cross 2	12128		176	50	198		37	21

Ratio majority recombinant to minority recombinant

Cross 1 (211)/(221) = 5.7
(121)/(221) = 5.7

Cross 2 (112)/(122) = 3.5
(112)/(212) = 4.75

Ratio minority recombinant to minority parent

Cross 2 (122)/(222) = 2.4
(212)/(222) = 1.75

majority recombinant/minority recombinant a_{112}/a_{122} can be calculated from these two crosses. In those cases in which we have more than one estimate of a class of recombinants the average is used. The values obtained are in good agreement with an estimate of five rounds of mating as can be seen by comparing the values found with those plotted in figure 4a.

(3) *Synchronous mating versus random-in-time mating.* In a cross with unequal input the minority parent is present as a very small fraction at genetic equilibrium, equal to the cube of the input fraction. In the case of synchronous mating this drop in frequency occurs fairly rapidly because every time a minority parent mates with a majority parent, there is only a chance of one in four of the minority parent not being eliminated by recombination. In the case of random-in-time mating, the elimination of the minority parent is not nearly as rapid, since there is an appreciable fraction of particles not participating in the mating, and this effect is even more pronounced if spread-in-time maturation is taken into account. These relations are illustrated in figure 3. For practical purposes it is more convenient to determine the ratio of minority recombinants to minority parent, rather than the absolute frequency of the minority parent, and this ratio is given in figure 4b. The minority recombinant serves as a measure of the number of bacteria actually infected with the minority parent. Figure 4b shows that for an input ratio of 9 to 1, and five rounds of matings, the ratio between the two genotypes is 5.7 in the case of synchronous mating, 2.6 in the case of random-in-time mating, and 1.4 if spread in maturation is taken into account.

Figure 4b is based on the following formulae, which are easily derived from the more general expressions given previously:

Synchronous mating

minor. recombinant

$$8a_{122}{}^{(n)} = (1 - 1/2^n)(1 - f^2)(1 - f + 2f/2^n) \qquad (13a)$$

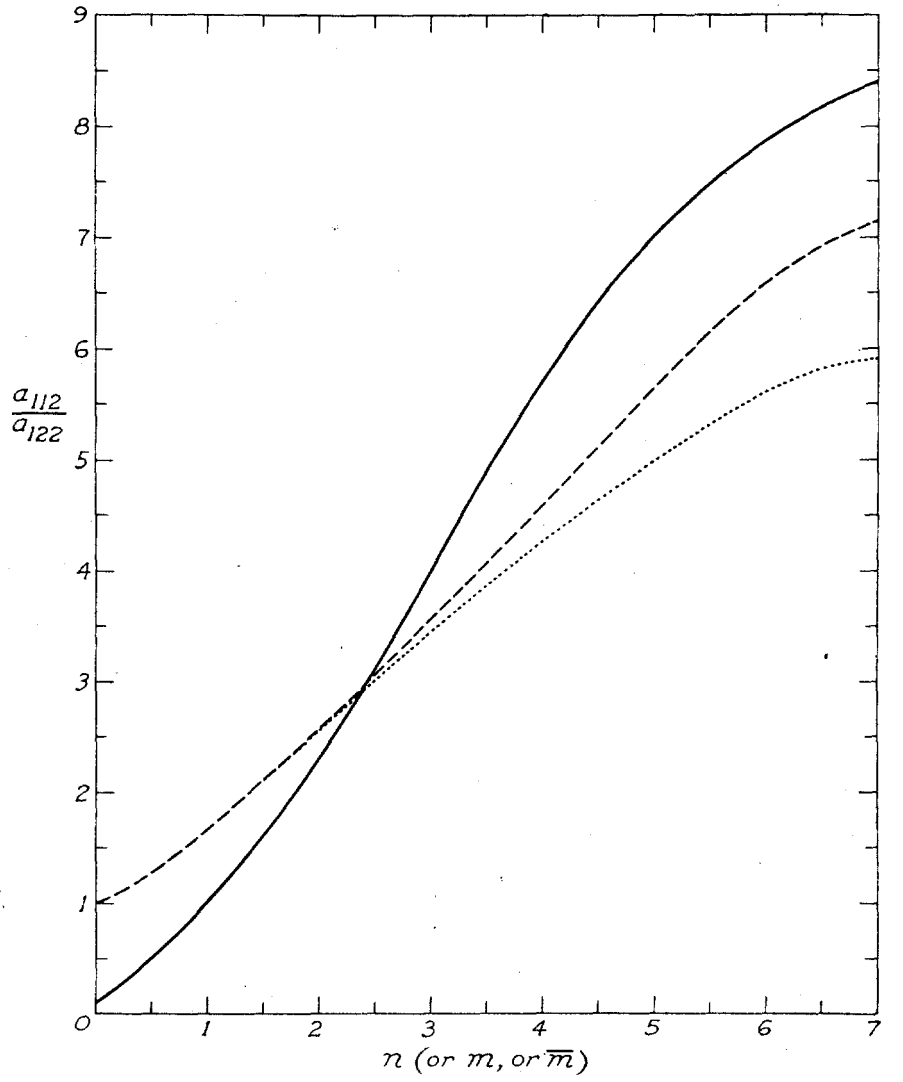

FIG. 4a.

minor. parent

$$8a_{222}^{(n)} = (1-f)^3 + (1-f^2)[3(1-f) + 2f/2^n]/2^n \qquad (13b)$$

Random-in-time mating

minor. recombinant

$$8a_{122}^{(n)} = (1-f^2)[(1-f(1-e^{-m/2}) + 2fe^{-m/2}(1-e^{-m/4})] \qquad (13c)$$

minor. parent

$$8a_{222}^{(n)} = (1-f)^3 + (1-f^2)[3(1-f)e^{-m/2} + 2fe^{-3m/4}] \qquad (13d)$$

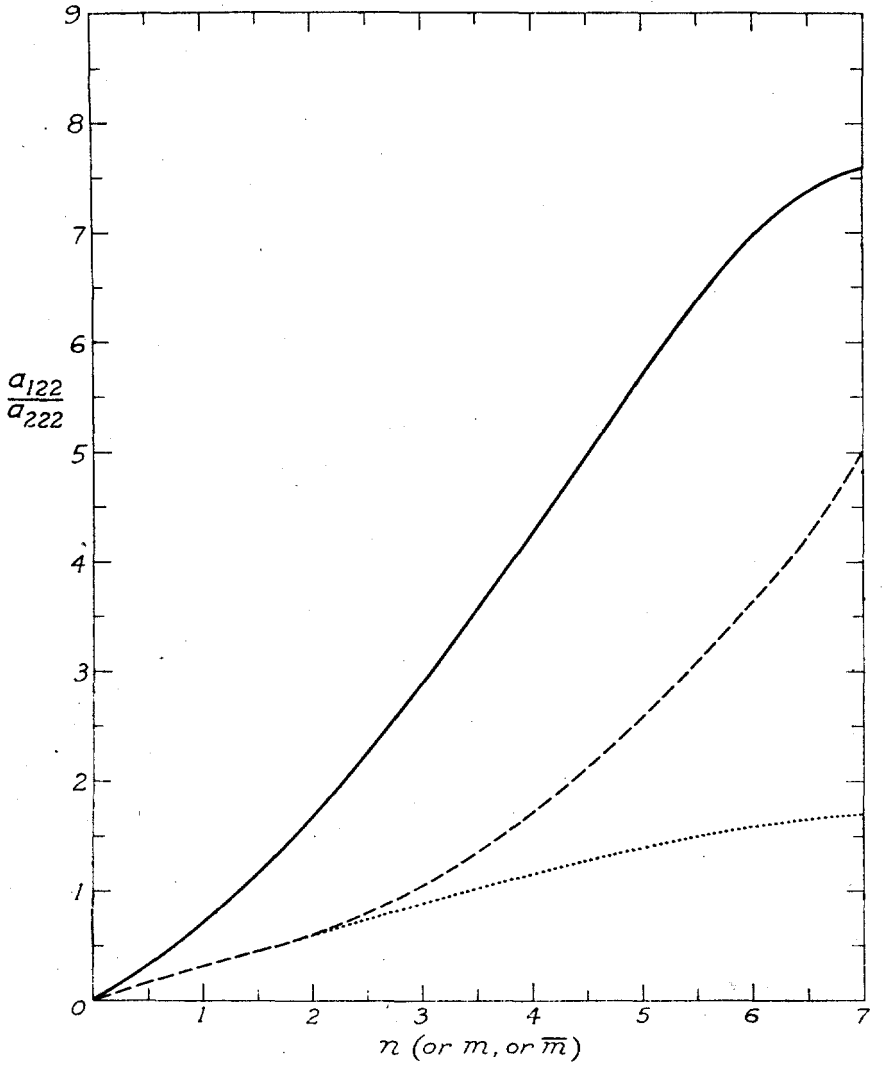

FIG. 4b.

FIGURE 4.—The ratio majority recombinant/minority recombinant (4a), and the ratio minority recombinant/minority parent (4b) in three-factor crosses, as a function of the number of rounds of mating. The continuous curves refer to synchronous mating (n rounds), the dashed curves to random-in-time mating (m rounds), the dotted curves incorporate the correction for the spread in maturation time (abscissa $\bar{m}$).

Only from cross 2 in table 5 can the ratio minority recombinant/minority parent (a_{122}/a_{222}) be calculated. In cross 1 the minority parent is pooled together with a minority recombinant as there is no way of distinguishing between *mr* and mr^+. The values found (2.4 and 1.75) are compatible, as can be seen from figure 4b, with the value of five rounds of mating, with or without the

correction for spread in maturation (2.6 and 1.4). Moreover this experimental value is in sharp disagreement with the value calculated for five rounds of matings assuming synchronous mating (5.7).

Two-factor crosses can also be used for this test, though the discrimination is not quite as sharp because the minority parent decreases only by a factor of 2 in every mating with the majority parent.

In figure 4c we represent, as a function of n (or m or $\bar{m}$), the ratio recombinant/minority parent for unequal input (f = .8) for synchronous mating and random-in-time mating. The calculations are based on equations 10a and 10b.

It will be seen that for five rounds of mating (n = 5) and a large excess of the majority parent (f = .8) the recombinants are more frequent than the genotype corresponding to the minority parent. This agrees with observation, as reported by DOERMANN and HILL (1953a). In table 6 the value recombin-

TABLE 6

Results obtained in the two crosses of table 5 considering only two markers, r and m. The majority parent (11) is hm^+ in cross 1 and h^+m in cross 2.

(*11*) × (*22*)

	(11)	(12)	(21)	(22)
Cross 1	12869	319	340	86
Cross 2	12128	226	198	58

Ratio recombinant to minority parent

Cross 1 = (12)/(22) = 4.0 (21)/(22) = 3.9

Cross 2 = (12)/(22) = 3.9 (21)/(22) = 3.4

ant/minority parent has been obtained from the data of table 5 using the three-factor cross as a two-factor cross by considering only the markers *r* and *m*. Comparing these ratios with those calculated in figure 2 we see that also in this case our data (3.4–4.0) are compatible with five rounds of random-in-time mating (4.7) and disagree with synchronous mating (6.8).

Linked factors

Having specified the parameters of the theory by the preceding experiments, we now wish to discuss the implications of the theory with regard to recombination between a pair of linked factors. The guiding idea of this analysis may be expressed as follows: in two-factor crosses, we may calculate an expression for the ratio $V_2 = a_{12}/a_{.2}$, i.e., the fraction in which a particular allele of the minority parent has recombined (equations 10a and 10b). We now wish to analyze how this fraction varies if, instead of considering the whole population, we confine our attention to that part of it in which the second factor has recombined with a third factor. Qualitatively, this amounts to the selection of a portion of the population which, on the average, has mated more often with unequal parents than the whole population. Furthermore, this selection will be the sharper the closer the linkage between the second and the third factor. We should expect, therefore, that the ratio $a_{121}/a_{.21}$ is greater than the ratio $a_{12}/a_{.2}$, and that the increase is more pronounced when the second and third

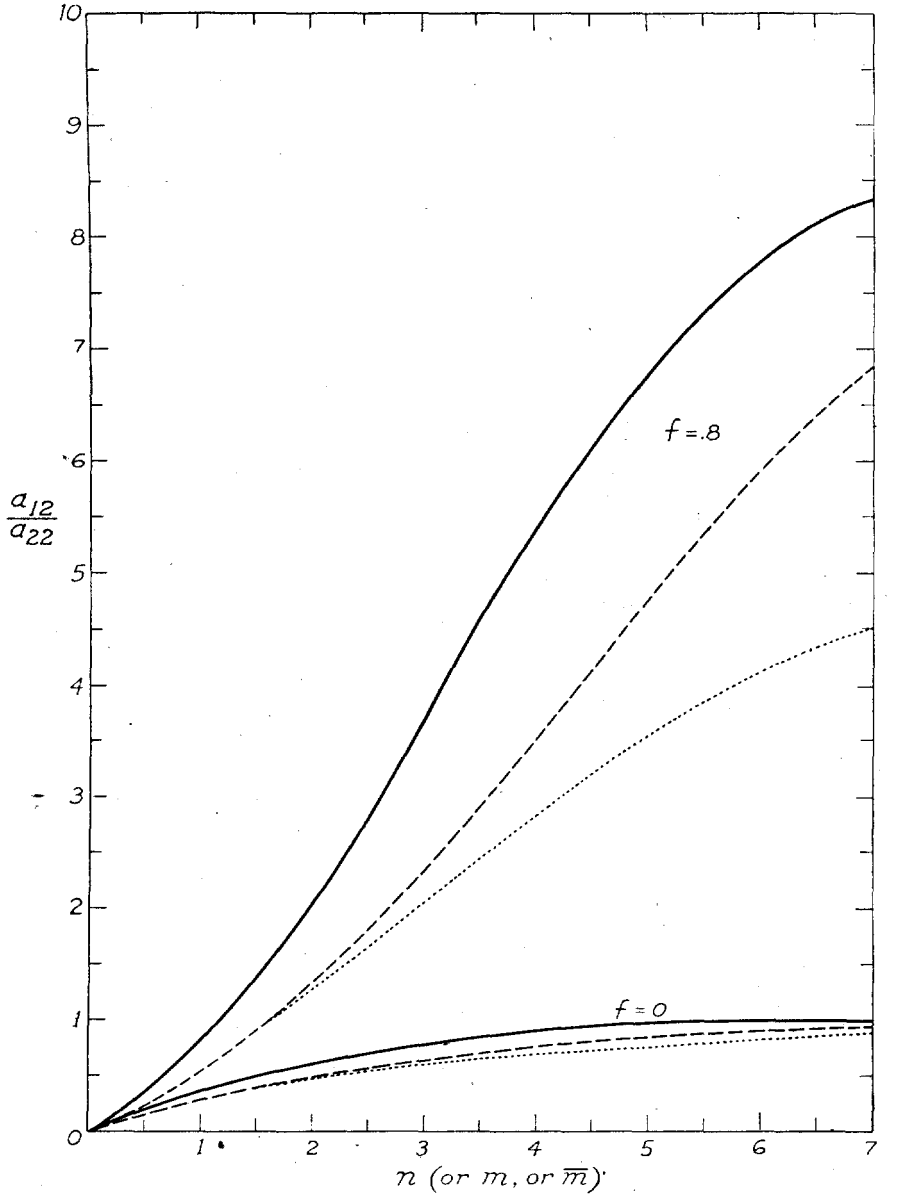

FIGURE 4c.—The ratio recombinant/minority parent (a_{12}/a_{22}) in two-factor crosses, as a function of the number of rounds of mating. The continuous curves refer to synchronous mating (n rounds), the dashed curves to random-in-time mating (m rounds), the dotted curves incorporate the correction for the spread in maturation time (abscissa $\bar{m}$). The upper curves are for unequal input, input ratio 9 to 1 ($f = .8$), the lower curves are for equal input ($f = 0$). Unlinked factors.

TABLE 7

V ratios, Synchronous Mating.

$$V_2 = a^{(n)}_{12}/a^{(n)}_{\cdot 2} = (1 + f)[1 - (1 - p)^n]/2$$
$$V_{ul} = a^{(n)}_{121}/a^{(n)}_{\cdot 21} = [1 + f(1 - 2/2^n)][1 - (1 - p)^n]/2(1 - 1/2^n)$$
$$V_{cl} = a^{(n)}_{121}/a^{(n)}_{\cdot 21} = [1 + f - (1 + f - 2p)(1 - p)^{n-1}]/2 \text{ (no interference)}$$
$$= (1 + f)[1 - (1 - p)^{n-1})]/2 \text{ (complete interference)}$$

V ratios, Random-in-time Mating.

$$V_2 = (1 + f)(1 - e^{-mp})/2$$
$$V_{ul} = [(1 + f)(1 - e^{-mp}) - 2fe^{-m/2}(1 - e^{-mp/2})]/2(1 - e^{-m/2})$$
$$V_{cl} = [1 + f - (1 + f - 2p)e^{-mp}]/2 \text{ (no interference)}$$
$$= (1 + f)(1 - e^{-mp})/2 \text{ (complete interference)}$$

factors are closely linked than when they are unlinked. We calculate the ratio $a_{121}/a_{\cdot 21}$ for the two limiting cases where the second and third factors are, respectively, unlinked, or very closely linked. These limiting values will be called V_{ul} and V_{cl}. In the formulas to be given p is the linkage parameter p_{12} for the first and second factor (equation 2). The linkage parameter $p_{23} = 1/2$ in case of V_{ul}. In the case of close linkage V_{cl} is independent of p_{23} in the first approximation. Table 7 gives the general formulas for any input ratio, for synchronous and random-in-time mating. In tables 8, 9, 10, 11 and 12 the results of crosses involving recombination between the locus *h* and the locus r_2 are shown. From these experiments we obtain estimates of the values V_2, V_{ul} and V_{cl} for values of f = 0 and of f = .8. By assuming an average value of 5 rounds of mating and random-in-time mating corrected for spread in maturation time, we first calculate (equation 10b) the value p from V_2 for f = 0, and obtain p = .14. Using this value p we calculate, using the formulas of table 7, all the other V values for f = 0 and f = .8. Table 13 shows the calculated and the experimental values.

Let us first consider the calculated values. Table 13 bears out the expectation that selection increases the V ratio, and that the increase is slight in the case of unlinked factors but appreciable in the case of linked factors. The selec-

TABLE 8

Fraction of recombinants (V_2) in crosses between h and r_2, with equal input (f = 0). Crosses 1 and 2 are the same cross made in coupling and repulsion. Cross 1 is h^+r_2 (11) × hr_2^+ (22). Cross 2 is hr_2 (11) × $h^+r_2^+$ (22).

(11) × (22)

	(11)	(22)	(12)	(21)
Cross 1	683	511	169	189
Cross 2	515	486	178	170

Ratio V_2 recombinants to total

Cross 1 $V_2 = (a_{12} + a_{21})/(a\cdot\cdot) = .23$
Cross 2 $V_2 = (a_{12} + a_{21})/(a\cdot\cdot) = .26$

TABLE 9

Fraction of recombinants V_{ul} between h and r_2 with equal input ($f = 0$). r_1 was chosen as third unlinked marker. Crosses 1 and 2 are the same cross made in coupling and repulsion. Each cross was repeated twice and each analyzed using the direct method (d.m.) and the sampling method (s.m.). Cross 1 is $h^+r_2^+r_1$ (111) × $hr_2r_1^+$ (222). Cross 2 is $hr_2^+r_1$ (111) × $h^+r_2r_1^+$ (222).

(111) × (222)

		(1 ••)	(2 ••)	(212)	(112)
Cross 1	d.m.	438	443	52	120
	s.m.			38	114
Cross 2	d.m.	1362	1395	174	399
	s.m.			42	110

Cross 1 d.m. V_{ul} = (212)/(• 12) = .30 *Cross 2* d.m. V_{ul} = (212)/(• 12) = .30
s.m. V_{ul} = (212)/(• 12) = .25 s.m. V_{ul} = (212)/(• 12) = .28

tive influence disappears almost completely if we assume strong interference between two exchanges within the same linkage group.

In addition, the calculations bring out three other features. First, that the increases are more pronounced in the case of random-in-time mating than in the case of synchronous mating; second, that the selective effect is greater in the case of equal input than in the case of very unequal input; third, that all the V ratios are almost twice as large in the case of very unequal input as in the case of equal input. The last point can be understood qualitatively as follows: the V ratios measure the fraction of the minority parents in which recombination between the first and second factor has taken place. This is twice as high in the case of very unequal input as compared to equal input because we assume that mating occurs at random with respect to partner. In the case of very unequal input, a particle of the genotype of the minority parent will nearly always mate with a particle of the genotype of the majority parent, while in the case of equal input, one half of the matings of a minority parental type will be with the same type. The increase in the V ratios when going from equal input to very unequal input is thus a direct consequence of the assump-

TABLE 10

Fraction of recombinants V_{cl} between h and r_2 with equal input ($f = 0$). r_7 was chosen as third marker closely linked to r_2. Crosses 1 and 2 are the same cross made in coupling and repulsion. Each cross was repeated twice and analyzed each time using the direct method (d.m.) and the sampling method (s.m.). Cross 1 is $hr_2^+r_7$ (111) × $h^+r_2r_7^+$ (222). Cross 2 is $h^+r_2^+r_7$ (111) × $hr_2r_7^+$ (222).

(111) × (222)

		(1 ••)	(2 ••)	(212)	(112)
Cross 1	d.m.	1585	1648	41	64
	s.m.			46	78
Cross 2	d.m.	773	836	21	34
	s.m.			71	118

Cross 1 d.m. V_{cl} = (212)/(• 12) = .39 *Cross 2* d.m. V_{cl} = (212)/(• 12) = .38
s.m. V_{cl} = (212)/(• 12) = .37 s.m. V_{cl} = (212)/(• 12) = .38

TABLE 11

Fraction of recombinants V_{ul} between h and r_2 with input 9 to 1 ($f = .8$). r_1 was chosen as third unlinked marker. Crosses 1 and 2 are the same cross made in coupling and repulsion. The frequency of the majority parent is indicated with the symbol (111). Cross 1 is $h^+r_2r_1^+$ (111) × $hr_2^+r_1$ (222). Cross 2 is $hr_2r_1^+$ (111) × $h^+r_2^+r_1$ (222). Both crosses were analyzed by the direct method only.

(111) × (222)

	(1··)	(2··)	(121)	(221)
Cross 1	11123	277	82	84
Cross 2	5112	168	50	43

Cross 1 $V_{ul} = (121)/(\cdot 21) = .50$ *Cross 2* $V_{ul} = (121)/(\cdot 21) = .54$

tion that mating is random with respect to partner, and its experimental verification may be considered as a test of this hypothesis.

The corrections due to spread in maturation time are in every case very small.

The experimental results shown in table 13 agree reasonably well with the theoretical predictions. Due to the fact that the experiments are very laborious, the number of plaques on which some of the estimates of genotype frequencies are based is rather small, as can be seen from the standard error of some of the values. Three groups of values are certainly different from each other. For f = 0, V_2 and V_{ul} are in the range of 24–29%, while V_{cl} is certainly greater (38%). For f = .8 all the V values are in the range 49–55%.

The experiments thus confirm some of the theoretical predictions and are not accurate enough to test certain others. They confirm:

(1) the increase of the V ratios in going from equal input to very unequal input, thus supporting the assumption that mating occurs between vegetative phages pairing at random.

(2) the selective influence of recombinations within the same linkage group, thus supporting the assumption that entire linkage groups are involved in the elementary acts of recombination, and suggesting that there is no strong interference between different exchanges within the same linkage group.

The experiments are not clear-cut in testing the selective influence of recombinations with an unlinked factor. The predicted effect is very small, the effect

TABLE 12

Frequency of recombinants V_{cl} between h and r_2 with input 9 to 1 ($f = .8$). r_7 was chosen as third marker closely linked to r_2. Crosses 1 and 2 are the same cross made in coupling and repulsion. The majority parent is indicated with the symbol (111). Cross 1 is $hr_2r_7^+$ (111) × $h^+r_2^+r_7$ (222). Cross 2 is $h^+r_2r_7^+$ (111) × $hr_2^+r_7$ (222). Both crosses were analyzed by the sampling method only.

(111) × (222)

	(121)	(221)
Cross 1	61	45
Cross 2	47	45

Cross 1 $V_{cl} = (121)/(\cdot 21) = .58$ *Cross 2* $V_{cl} = (121)/(\cdot 21) = .50$

found experimentally is in the same direction and slightly larger than the predicted value, but the standard errors are a little too large for this to be convincing. Within the group of experiments here reported, this is the only fairly direct test of the assumption that in the elementary act of recombination not only whole linkage groups but also entire vegetative phages are involved. It would be quite cumbersome to do this experiment on such a scale as to verify or exclude with certainty the small theoretically predicted difference between V_{ul} and V_2. The point in question could be tested more efficiently in four-factor crosses with two linked factors on each of two linkage groups and determining the fraction of recombinants for one pair of linked factors among that portion of the progeny which carries the recombination for the other pair

TABLE 13

Fraction of recombinants (V) between loci h and r_2.

	Experimental	Calculated			
		(1)	(2)	(3)	(4)
f = 0					
$V_2\ a_{12}/a_{.2}$	24.5 ± .0078	.25	....	.26	.24
$V_{ul}\ a_{121}/a_{.21}$	.29 ± .014	.27	....	.27	.27
$V_{cl}\ a_{121}/a_{.21}$	.38 ± .022	.32	.25	.30	.31
f = .8					
$V_2\ a_{12}/a_{.2}$	.49 ± .07	.45	....	.48	.43
$V_{ul}\ a_{121}/a_{.21}$	.51 ± .031	.47	...	.43	.47
$V_{cl}\ a_{121}/a_{.21}$	.54 ± .041	.52	.45	.48	.50

The V ratios have been calculated on the following assumptions:

(1) Random-in-time mating, no interference.
(2) Random-in-time mating, complete interference.
(3) Synchronous mating, no interference.
(4) Random-in-time mating, no interference, corrected for spread in maturation time.

The number of matings is estimated to be 5 from the crosses with unlinked characters. The linkage parameter p is estimated from the value of V_2 for f = 0, which gives p = .14.

of linked factors. Suitable markers for such an experiment are not yet available in T2.

In the case of unequal multiplicity, the theory predicts a lesser shift in the V ratios due to selective influences, than for equal input. In this respect, the theory is confirmed by the experiments with fair certainty. The actual value of the shift is also of the right order, and in the right direction, but the standard errors are too large to consider this point of the theory confirmed.

DISCUSSION

In the foregoing pages we have presented a theory of genetic recombination in phage according to which a mixed infection of a bacterium with different genetically marked types represents not a simple genetic cross but an experi-

ment in population genetics. The intracellular population of vegetative phages undergoes several rounds of mating. Mating is random with respect to partners and, presumably, random with respect to time. Superimposed on these processes is the maturation of the vegetative phages, which removes them from the pool of multiplying and mating particles, and which is assumed to be random with respect to genotype.

The general purpose of the theory is to explain and predict the results of crosses (mixed infections). The variables of the experiments are the markers and their linkage relations, the ratio between the infecting parental particles, and the time of inspection (at normal lysis time, or earlier, in the case of premature lysis, or later, in the case of delayed lysis).

The theory fits with some very general notions concerning the life cycle of bacteriophages, namely, the transformation into vegetative phage and the accumulation at a linear rate of mature phages from a pool of growing and recombining vegetative phages, as derived both from genetic and from biochemical tracer studies.

The principal accomplishment of the theory is that it accounts for (1) the occurrence of particles in the progeny which combine markers from more than two parents, (2) the deviations from genetic equilibrium and their decrease in the course of time, and (3) the positive correlation between different recombinations in the progeny population. This positive correlation should and does (DOERMANN and HILL 1953b) also show up in simple mapping experiments with two-factor crosses. Map distances (i.e., the linkage parameters p) calculated by our equation (10b) show neither positive nor negative correlation, suggesting that there is little, if any, interference between different exchanges within the same linkage group.

Our theory may be contrasted with an hypothesis suggested some time ago by STURTEVANT (cf. HERSHEY and ROTMAN 1949) regarding the origin of recombinants: this hypothesis assumes that the progeny phage particles are copied from patterns provided by the parental particles. Half finished copies can switch from one pattern to another, thus producing recombinants. A formal development of this hypothesis requires additional specifications, of which the most important is a specification as to whether or not copies, including unfinished copies, in turn become patterns. We will assume that the copies do *not* become patterns since otherwise the process would lead to a very rapid progressive fragmentation of the genetic material. In a formal development assuming *fixed* patterns the switching from one pattern to another would be equivalent to one mating in our theory. The principal characteristic of this hypothesis is that reciprocal recombinants are made in statistically independent acts. It thus accounts elegantly for the lack of correlation between reciprocal recombinants observed by HERSHEY and ROTMAN (1949). Our theory also accounts for this lack of correlation, through the randomizing influences of multiplication and maturation after recombination. On the other hand, this hypothesis permits the occurrence of only one recombination per switching, which is equivalent to one mating. It implies, therefore, complete interference in any one mating for different recombinations within the same linkage group, and this, as we have

shown, is in conflict with the findings here reported and with those of DOERMANN and HILL (1953b). Another implication of STURTEVANT's hypothesis is that the genetic composition of the intracellular phage population should be independent of time, and this, too, is contradicted, both by experiments with premature lysis (DOERMANN and HILL 1953a) and with delayed lysis (LEVINTHAL and VISCONTI 1953).

One aspect of phage genetics which has not been accounted for by our theory is the work of HERSHEY and CHASE (1951) on heterozygous phage particles in the progeny. HERSHEY finds that about 2% of the progeny from a bacterium infected with equal input of two marked parents is heterozygous with respect to any given factor. At first sight one is tempted to interpret these particles by the assumption that a certain random fraction of the zygotes of vegetative particles in the mating pool mature without segregating. In two-factor crosses this leads to the prediction that heterozygosis with respect to one of the two factors should be strongly correlated with heterozygosis with respect to the other factor, and that this correlation should increase with the degree of linkage between the two factors. In contradiction with this HERSHEY finds zero correlation when the two factors are unlinked. When the factors are closely linked HERSHEY does find a correlation but not nearly as large a one as required by our assumptions. We see no simple way of accounting for this discrepancy and therefore consider it unlikely that the heterozygotes should be interpreted as a sample from the zygotes in the mating pool.

We may conclude by saying that, from the point of view of this theory, mixed infection constitutes a very complex genetic experiment, and that the complications are mainly due to the experimental difficulties which stand in the way of an analysis of single matings. Particularly, the idea that in the elementary mating process reciprocal recombinants are produced in exactly equal numbers seems at present not susceptible to any direct experimental tests. On the whole, our theory tends to place the genetics of phage very much in line with orthodox genetical theory, as it has been worked out for higher organisms.

SUMMARY

A theory of recombination of genetic markers of phage particles during mixed infections of bacteria is developed. The theory assumes that the parental particles upon entering the host cell become vegetative particles, that these particles multiply, mate pairwise, repeatedly, and random with respect to partner. Further, that maturation of vegetative particles into mature infective particles is irreversible within the mother cell, and that the matured particles do not multiply and do not mate.

The population genetics problem thus posed is solved quite generally for two- and three-factor cases and worked out in detail for several special cases.

Experimental results for several two- and three-factor crosses, involving either unlinked factors, or factors linked in various ways, are presented and are compared with the theoretical predictions. These results are in full agreement with the theory if the average number of matings per vegetative particle is taken to be five. The experiments may be considered as supporting specifi-

cally the assumption that mating is random with respect to partner, and that at least whole linkage groups are involved in the elementary act of recombination. Further, interference between different recombinations within the same linkage group, if it occurs at all, is far from complete.

A three-factor test is described for which certain genotype ratios in the case of nonlinkage are exactly 1 : 1, independent of various influences which lower this ratio in two-factor crosses in an amount that cannot well be predicted. Experimental results are presented which confirm this 1 : 1 ratio.

ACKNOWLEDGMENTS

The authors wish to record their indebtedness to DR. N. SYMONDS, PROFESSOR H. F. BOHNENBLUST, and MR. DALE KAISER, for their help in clarifying the algebraic aspects of the problem; to DR. A. D. HERSHEY, for providing the genetically marked stocks and for invaluable advice concerning the execution of the experiments; to DR. A. H. DOERMANN, for many discussions; and to MISS BARBARA POWELL and to MR. ROBERT COON for technical assistance.

APPENDIX

A theorem on random-in-time mating (N. Symonds)

The effect of random-in-time mating for a period during which each particle mates on the average m times is equal to that of splitting the total population into non-intermating groups of size $p_1, p_2, \ldots$, where $p_r(m) = m^r e^{-m}/r!$, and letting each group undergo r rounds of synchronous mating.

The proof of this theorem rests on the fact that equation (6a), which expresses the relation between the genotype frequencies in one generation in terms of those in the preceding generation and which appears to be a quadratic relation is actually linear because the sums of genotype frequencies represented by the symbols involving two dots are not variables but constants, since they represent allele frequencies.

Equation (6a) may thus be written symbolically as

$$a^{(n)} = La^{(n-1)} \tag{1}$$

where a represents the set of eight numbers a_{ijk} and L represents a linear operator.

When mating is random-in-time, the gene frequencies become functions of time rather than of the number of rounds of mating n, and the recurrence relation (6a) is replaced by a differential equation. If we choose as the time unit the time in which a particle mates on the average once then in the time element dm the fraction a dm mates and is replaced by the fraction L a dm. Thus the differential equation for a is

$$da/dm = La - a \tag{2}$$

or

$$d(ae^m)/dm = Lae^m \tag{2a}$$

Symonds' theorem may be stated by saying that the solution of this differential equation which satisfies the initial condition is

$$a = \sum_r p_r \, a^{(r)} \tag{3}$$

where

$$a^{(r)} = L^r a^{(0)}$$

is the distribution of genotypes after r rounds of synchronous mating. To verify that (3) is the desired solution of (2) we observe, first, that for $m = 0$ we obtain from (3)

$$a(0) = a^{(0)}$$

since, for $m = 0$, $p_0 = 1$ and all other $p_r = 0$. Second, substituting (3) into (2a) we obtain on the left

$$d(ae^m)/dm = \Sigma a^{(r)}\ d(p_r e^m)/dm = \Sigma a^{(r)}\ p_r\ e^m\ r/m \qquad (4)$$

and on the right

$$Lae^m = L\Sigma p_r a^{(r)} e^m = \Sigma a^{(r+1)} p_r e^m \qquad (5)$$

The two expressions (4) and (5) are equal, as may be seen by comparing the coefficients of the same $a^{(r)}$ in both series and noting that the fractions p_r obey the recurrence relation

$$p_r\ r/m = p_{r-1}$$

The crucial step in this proof occurs in equation (5) where we make use of the fact that the operator L operating on a sum of distributions is equal to the sum of the distributions obtained by letting the operator operate on each term of the sum separately. It is here that we make use of the linearity of the operator. As may be seen by reference to the more general treatment given by GEIRINGER (1944), it is only in one-, two-, and three-factor crosses that the recurrence formula corresponding to (6a) can be looked upon as a linear relation. When more genetic factors are involved products occur in the recurrence formula of which neither factor is a constant. Then the recurrence relation is truly quadratic and the solution of the differential equation (2) is not of the simple form (3).

LITERATURE CITED

BENZER, S., 1952 Resistance to ultraviolet light as an index to reproduction in bacteriophage. J. Bact. **63:** 59–72.

DOERMANN, A. H., 1948 Intracellular growth of bacteriophage. Carnegie Institution of Washington Yearbook **47:** 176–186.

DOERMANN, A. H., and M. B. HILL, 1953a The intracellular growth of bacteriophage. III. Intracellular aspects of genetic recombination. Genetics (in press).

DOERMANN, A. H., and M. B. HILL, 1953b Genetic structure of bacteriophage T4 as described by recombination studies of factors influencing plaque morphology. Genetics (in press).

DULBECCO, R., 1949 The number of particles of bacteriophage T2 that can participate in intracellular growth. Genetics **34:** 126–132.

1952a A critical test of the recombination theory of multiplicity reactivation. J. Bact. **63:** 199–207.

1952b Mutual exclusion between related phages. J. Bact. **63:** 209–217.

GEIRINGER, H., 1944 On the probability theory of linkage in Mendelian heredity. Ann. Math. Statistics **15:** 25–57.

HERSHEY, A. D., and R. ROTMAN, 1948 Linkage among genes controlling inhibition of lysis in a bacterial virus. Proc. Nat. Acad. Sci. **34:** 89–96.

1949 Genetic recombination between host-range and plaque-type mutants of bacteriophage in single bacterial cells. Genetics **34:** 44–71.

HERSHEY, A. D., and MARTHA CHASE, 1951 Genetic recombination and heterozygosis in bacteriophage. Cold Spring Harbor Symp. Quant. Biol. **16:** 471–479.

LEVINTHAL, C., and N. VISCONTI, 1953 Recombination and reproduction in bacteriophage (in press).

LURIA, S. E., 1947 Reactivation of irradiated bacteriophage by transfer of self-reproducing units. Proc. Nat. Acad. Sci. **33:** 253–264.

LURIA, S. E., and R. DULBECCO, 1949 Genetic recombination leading to production of active bacteriophage from ultraviolet inactivated bacteriophage particles. Genetics **34:** 93–125.