Part One:

SOME THEORETICAL PROBLEMS ARISING IN THE GENETICS OF BACTERIOPHAGE

Part Two:

A CRITICAL TEST OF A CURRENT THEORY OF GENETIC RECOMBINATION IN BACTERIOPHAGE

Thesis by

Charles Myron Steinberg

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1961

ACKNOWLEDGEMENTS

I am deeply indebted to:

MAX DELBRÜCK--for patience, understanding, encouragement, and guidance.

ROBERT EDGAR and FRANK STAHL--for active scientific collaboration.

RICHARD EPSTEIN and JEAN WEIGLE--for their interest in and encouragement of the work described herein.

LOIS GLASS EDGAR and DON RIDGEWAY--for invaluable help in the preparation of the manuscript.

ILGA LIELAUSIS--for technical and other assistance.

THE NATIONAL SCIENCE FOUNDATION, E. I. DU PONT DE NEMOURS AND COMPANY, THE NATIONAL FOUNDATION, and THE U. S. PUBLIC HEALTH SERVICE--for financial aid.

ABSTRACT OF PART ONE

Calculations are presented giving the distribution of clone-sizes of mutants arising in an "equilibrium pool" of vegetative phage. It is assumed that replication and maturation are random-in-time processes. The equilibrium pool model is an idealization of current notions about a major portion of the phage life cycle. The calculations are compatible with experimental measurements of the mutant clone-size distribution. It is also shown that the model can account for the observed variance in total burst size.

A generalized theory of the kinetics of genetic recombination in phage is presented. Essential assumptions of the theory are that recombination occurs as a result of discrete interactions or matings, and that the particles which participate in a mating are randomly chosen from the entire population of the vegetative pool. No restrictions are placed upon the number of particles that participate in the mating or upon the mechanism by which recombination occurs as a result of the mating. Experimental observations leading to the theory are discussed as well as the prospects for experimental determinations of the participation number and the mechanism of recombination.

ABSTRACT OF PART TWO

Results are reported concerning the segregation pattern of loosely linked outside markers within a selected class of particles which are recombinant for closely linked internal markers:

(a) The outside markers maintain qualitatively normal linkage relations with the internal markers and with each other. This is true whether one or two exchanges are required to produce the selected recombinant class.

(b) The region of HNI extends into the intervals adjacent to those in which the exchanges are selected but does not extend into a non-adjacent interval.

(c) Multiple exchanges within the selected class occur with little, if any, additional interference.

(d) A significant fraction of the selected recombinant class is heterozygous for at least the adjacent outside markers.

In summary, aside from the "complications" of negative interference and heterozygosis, the segregation pattern of outside markers is "normal."

In view of current notions concerning the mechanism of recombination in phage, the "complications" are to be expected, while the "normal" result leads to a paradox. According to observations of other workers, recombinants for very closely linked markers arise as segregation products of phage heterozygotes, and furthermore, such heterozygotes are recombinant for outside markers. On the other hand, result (a) above means that recombinants for closely linked markers are not necessarily recombinant for outside markers. Some possible resolutions of this paradox are discussed.

TABLE OF CONTENTS

PART ONE:

SOME THEORETICAL PROBLEMS ARISING IN THE GENETICS OF BACTERIOPHAGE

PART ONE

I. THE CLONE SIZE DISTRIBUTION OF MUTANTS ARISING FROM A STEADY-STATE POOL OF VEGETATIVE PHAGE.

1. Introduction.

In 1951, Luria (1) studied the frequency distribution of the number of spontaneous phage mutants arising in single cell bursts. His results were in very satisfactory agreement with the following model: one intracellular phage "gene" gives rise to two genes, each of which in turn gives rise to two genes, and so on. This we will refer to as exponential growth. He further assumed that the replication process is completely synchronized so that all lines of descent are strictly equal. It is easily shown that such a regime leads to the probability, y_n, that a mutant clone consists of n individuals, which is given by

$$y_n = 1/n, \text{ if n is an integral power of 2}$$

$$= 0, \text{ otherwise}$$

But replication is not completely synchronized and clone sizes other than integral powers of two are observed. Luria circumvented this difficulty by using the accumulated probability, Y_n, that a clone contains n or more mutants, which for integral powers of two is given by

$$Y_n = 2/n$$

A quantitative comparison of experiment with theory is then made by plotting log Y_n *vs.* log n, the prediction for an exponential growth process being a linear relationship with slope - 1. Luria's scheme is, of course, just an idealized version of the growth habit of bacteria, for which the theoretical mutant distribution has been greatly elaborated by many authors since the problem was introduced by Luria and Delbrück in 1943 (2). (For later references see (3).)

The currently accepted notion of phage growth differs somewhat from the above (4). A mature phage is an inert particle, which, in particular, shows no measurable rate of mutation or genetic recombination. After adsorption to a sensitive host cell, the genetic determinants of the phage are injected into the cell. The phage is then in the so-called vegetative state; it is non-infectious but endowed with the ability to replicate and undergo genetic recombination. After some delay, the number of vegetative phages increases exponentially for a time. Then vegetative phage begin to be removed from the "pool" and matured. Such mature phage are inert, like their parents, and do not participate further in the growth processes of the vegetative phage pool. The rate of maturation balances the rate of replication, resulting in a steady-state pool of vegetative phage which is of constant size.

Thus it is seen that current notions about phage growth involve exponential replication, but with markedly unequal lines of descent. It might be intuitively felt, perhaps, that formally this is not too different from the bacterial scheme. Furthermore, the good agreement of the experimental results with the very simple theory did not prompt the consideration of more complicated alternatives. So the matter stood until 1956 when F. W. Stahl and G. Streisinger definitely confirmed the tentative conclusion of Stahl (5) that the distribution of rare phage recombinants in single cell bursts is quite different from the distribution of mutants. According to ideas current at that time, there was no reason why rare recombinational events should give rise to a different clone size distribution than rare mutational events. Therefore, Stahl and Streisinger wondered whether their results might not be what is to be expected from steady-state pool kinetics (which would, of course, then call for a special explanation of Luria's results). This was the motivation for the present study. We have written the solution for the case of a vegetative pool of statistically constant size where maturation and multiplication are concurrent random-in-time processes. Levinthal (6) has considered a similar case. He supposes, for convenience, that the pool undergoes alternating doublings by multiplication and halvings by maturation. Our analysis is probably more realistic (7), certainly more elegant, and, in

any event, emphasizes the effect of unequal lineages to the greatest degree possible.

2. Mathematical Exposition.

Let us first describe the concepts introduced in the preceding section in a somewhat more explicit and precise fashion. We are concerned with a collection of elements or "particles" which may be classified in two ways--by "genotype" and by "state." With respect to genotype, a particle may be "wild-type" or "mutant"; with respect to state, a particle may be "vegetative" or "mature." Thus, there are four distinguishable types of particles. The state of the system as a whole is specified by the number of particles of each type. A vegetative particle of a given genotype may undergo three types of transitions--"replication," "maturation," and "mutation." Replication is a transition from one vegetative particle of a given genotype to two particles of the given genotype. Maturation is a transition from the vegetative to the mature state without change of genotype. Mutation is a transition from one genotype to the other without change of state. A mature particle undergoes no transitions whatsoever. In these terms, our problem is the following: given N vegetative, wild-type particles at time zero, what is the probability, $\overline{P}_n(T)$, that there are exactly n mature, mutant particles at time T? An explicit solution of this problem

requires further specifications concerning the transition probabilities.

We postulate that replication and maturation are independent Poisson processes (8) with identical rate constants. That is, if p(t) is the probability that a particle replicates (or matures) in time t, then

$$\lim_{t \to 0} p(t) / t = k \qquad (1)$$

where k is a constant independent of the time of the last replication. Since the parameter, k, is the same for maturation and replication, it is evident that the average number of vegetative particles is a constant, N. For convenience, we will choose a time scale such that k=1.

Concerning the mutational transition, the only biologically interesting cases are those in which the probability is extremely small. Accordingly, we will restrict ourselves to the limiting distribution. Specifically, we assume that one and only one mutational transition occurs in the time interval (0,T). Furthermore, since the average number of vegetative particles is constant, it follows that the time of the mutational transition is uniformly distributed over (0,T).

Let $P_n(x)$ = the probability that a mutant, vegetative particle present at time zero will give rise to n mature mutant particles at time x.

$q_n(x)$ = the probability that a mutant, vegetative particle present at time zero will give rise to n vegetative mutant particles at time x.

First we consider a time interval (0,h) and assume, for the moment, that the mutation has occurred at time zero. When h is small, we need consider only the following three cases: the mutant particle matures (probability q_0); the mutant particle replicates once (probability q_2); the mutant particle neither replicates nor matures (probability q_1). It follows from equation 1 that the probabilities for more complex events (involving more than one transition) will be second order or higher in h and would vanish in the limit we will presently take. It also follows directly from equation 1 (and recalling the convention that k=1) that the relevant probabilities, correct to the first order in h, are:

$$\begin{aligned} q_0(h) &= h \\ q_1(h) &= 1-2h \\ q_2(h) &= h \end{aligned} \qquad (2)$$

We now consider an adjacent time interval (h,t+h) where t is not necessarily small. At time h there are 0, 1, or 2 mutant vegetative particles. Since these are mutually exclusive possibilities, we can write an expression for $P_n(t+h)$ as a sum of three terms: each term consists of a q, say q_r, multiplied by the conditional prob-

ability that given r vegetative mutant particles at time h, there will be n mature particles at time t+h. The conditional probabilities can be expressed in terms of the $P_n(t)$ as follows:

(a) Coefficient of $q_0(h)$. If there are no mutant vegetative particles at time h, the one present at time zero must have matured. Since mature particles undergo no transitions, there will be one and only one mature mutant particle at any later time. Hence the coefficient of q_0 will be zero in all $P_n(t+h)$ except $P_1(t+h)$, in which case it will be unity.

(b) Coefficient of $q_1(h)$. If there is one mutant vegetative particle at the beginning of the interval (h,t+h) (there can be no mature mutant particles present), the conditional probability that n mature mutant particles will be present at the end of this interval is by definition $P_n(t)$.

(c) Coefficient of $q_2(h)$. If there are two mutant vegetative particles at the beginning of the interval (h,t+h), then at the end of this time interval one particle will have given rise to, for instance, i mature mutant particles and the other will have given rise to j mature mutant particles. The conditional probability of this event is the product $P_i(t)\ P_j(t)$. The conditional probability that there will be exactly n mature mutant particles at the end of the interval (h,t+h) is the sum

of all such products for which i+j=n. Accordingly we may write:

$$P_0(t+h) = q_1(h)P_0(t) + q_2(h)P_0^2(t)$$

$$P_1(t+h) = q_0(h)+q_1(h)P_1(t)+q_2(h)[P_0(t)P_1(t)+P_1(t)P_0(t)] \qquad (3)$$

$$P_2(t+h) = q_1(h)P_2(t)+q_2(h)[P_0(t)P_2(t)+P_1^2(t)+P_2(t)P_0(t)]$$

. . . . . .

$$P_n(t+h) = q_1(h)P_n(t) + q_2(t)[P^*P]_n \quad , \quad n \neq 1$$

where $[P^*P]_n$ is the n'th component of the convolution of the distribution of P_n with itself:

$$[P^*P]_n = P_0P_n + P_1P_{n-1} + P_2P_{n-2} + \cdots + P_{n-1}P_1 + P_nP_0$$

Substituting (2) into (3), rearranging, and dividing by h, we get in the limit $h \longrightarrow 0$:

$$\begin{aligned} P_0' &= -2P_0 + P_0^2 \\ P_1' &= 1 - 2P_1 + [P_0P_1 + P_1P_0] \\ P_2' &= -2P_2 + [P_0P_2 + P_1^2 + P_2P_0] \\ &\cdots \\ P_n' &= -2P_n + [P^*P]_n \quad , \quad n \neq 1 \end{aligned} \qquad (4)$$

where $$P_n' = \frac{dP_n(t)}{dt}$$

In principle, the problem is now solved. The differential equations 4 could be solved successively for the P_n's. The labor involved, however, leads us to seek an easier method. Therefore, we define the generating function of the P_n as

$$F(s,t) = P_0 + P_1 s + P_2 s^2 + \cdots = \sum_{n=0}^{\infty} P_n s^n \tag{5}$$

Squaring F and collecting the coefficients of each power of s, we have that

$$F^2(s,t) = [P^*P]_0 + [P^*P]_1 s + [P^*P]_2 s^2 + \cdots = \sum_{n=0}^{\infty} [P^*P]_n s^n \tag{6}$$

If we multiply the n'th equation 4 by s^n and then sum over all the equations we obtain

$$\sum_{n=0}^{\infty} P_n' s^n = s - 2 \sum_{n=0}^{\infty} P_n s^n + \sum_{n=0}^{\infty} [P^*P]_n s^n \tag{7}$$

which is equivalent to

$$F' = s - 2F + F^2 \tag{8}$$

where

$$F' = \frac{\partial F(s,t)}{\partial t}$$

The solution of this differential equation with the initial condition that $P_0 = 1$ at $t = 0$ is

$$F = 1 - (1-s)^{1/2} \tanh(1 - s)^{1/2} t \tag{9}$$

Thus far we have assumed that the mutational transition occurred at time t=0. To find the clone size probabilities for the case that the probability of mutation is uniformly distributed over the interval from t=0 to t=T, we average over this time interval, giving equal weight to all times. That is, the desired probabilities are

$$\overline{P}_n = \frac{1}{T}\int_0^T P_n(t)\,dt \tag{10}$$

and the corresponding generating function is

$$G(s,T) = \frac{1}{T}\int_0^T F(s,t)\,dt = 1 - \frac{1}{T}\log_e \cosh(1-s)^{1/2} T \tag{11}$$

since

$$G = \overline{P}_0 + \overline{P}_1 s + \overline{P}_2 s^2 + \cdots + \overline{P}_n s^n + \cdots \tag{12}$$

Equation 11 is the generating function, in closed form, for the desired probabilities $\overline{P}_n$. We can recover a particular $\overline{P}_n$ by extracting the coefficient of s^n by the usual procedure for a power series expansion. That is, we evaluate the n'th derivative of G with respect to s at s=0 and divide by n!:

$$\overline{P}_n = \frac{1}{n!}\left(\frac{\partial^n G}{\partial s^n}\right)_{s=0} \tag{13}$$

These derivatives, which are easily taken for small values

of n, become prohibitively difficult for n greater than about 4. Fortunately, it is easy to express the derivatives as an infinite series suitable for numerical calculation. Cosh x may be expanded as an infinite product (9):

$$\cosh x = \prod_{i=0}^{\infty} \left[1 + \frac{x^2}{(i + 1/2)^2 \pi^2} \right] \tag{14}$$

Therefore we can rewrite equation 11 as

$$G(s,T) = 1 - \frac{1}{T} \sum_{i=0}^{\infty} \log_e \left[1 + \frac{(1 - s)T^2}{(i + 1/2)^2 \pi^2} \right] \tag{15}$$

Taking the n'th derivative, dividing by n!, and evaluating at s=0, we arrive at

$$\overline{P}_n = \frac{(2T)^{2n}}{nT} \sum_{i=0}^{\infty} \left[\frac{\pi^2}{(2T)^2 + (2i + 1)^2 \pi^2} \right]^n \tag{16}$$

This series converges quite rapidly for n greater than 3. For $T \geq 2\pi$, it may be shown that for n=3, 7 terms give an error of less than 0.1%; for n=4, 5 terms give 0.1% accuracy; for n=5-7, 4 terms suffice; while only 3 terms are needed to give the same accuracy when n is 8 or higher.

Equation 16 gives the desired probabilities in terms of a single temporal parameter, T. We can evaluate this parameter independently of the probabilities, $\overline{P}_n$. To do this, we must consider another problem: given N vegetative particles at time zero, what is the probability, $R_B(T)$

that there are exactly B mature particles at time T? That is, we seek the distribution of mature particles, without respect to genotype. We have already solved this problem for the special case of N=1. In this case $R_B(T)=P_n(t)$, with B=n, and the generating function, F, is given by equation 9. In the argument leading to equation 3, it was shown that for the case N=2, $R_B=[P^*P]_n$, with B=n. It was then shown that $[P^*P]_n$ is generated by the function, F^2. It is easy to generalize this result by induction. We define the generating function of the R_B as

$$H(s,T) = R_0+R_1s + R_2s^2 + \cdots \tag{17}$$

Then

$$H(s,T) = F^N(s,T) \tag{18}$$

where F is given by equation 9.

It is, in principle, possible to calculate the entire set of probabilities $R_B(T)$ from equation 18 by methods similar to those used previously. For our purposes, however, it suffices to calculate the mean and variance of the distribution. These moments can be calculated directly from the generating function without explicit evaluation of the separate probabilities. This is achieved by differentiating H with respect to s and evaluating the derivatives at s=1. In particular we have the mean

$$\overline{B} = \left(\frac{\partial H}{\partial s}\right)_{s=1} \tag{19}$$

and the variance

$$\mathrm{Var}(B) = \left(\frac{\partial^2 H}{\partial s^2}\right)_{s=1} - \bar{B}\,(\bar{B} - 1) \tag{20}$$

In terms of the function, F, these expressions are:

$$\bar{B} = N\left(\frac{\partial F}{\partial s}\right)_{s=1} \tag{21}$$

$$\mathrm{Var}\,(B) + N\left(\frac{\partial^2 F}{\partial s^2}\right)_{s=1} - \frac{\bar{B}^2}{N} + \bar{B} \tag{22}$$

Evaluating the derivatives, we arrive at the following expressions:

$$\bar{B} = N \cdot T \tag{23}$$

$$\mathrm{Var}\,(B) = \bar{B}(2/3T^2 - T + 1) \tag{24}$$

Both $\bar{B}$ and N (or $\bar{N}$, if N is not a constant) are, in principle, measurable quantities, and T may be evaluated from these using equation 23. We shall use the variance relation (equation 24) in the discussion which follows.

3. Results and Discussion

Our model predicts that a certain proportion of all mutational events will fail to give rise to mature mutant particles. Such events are, of course, not experimentally observable. Accordingly, the calculated probabilities $\bar{P}_n$ must be renormalized by dividing by $1 - \bar{P}_0$ for comparison

with experimental data. Conversely, the experimental data would need to be corrected if it were desired to calculate absolute mutation rates. In any event, the correction is small (about 10% for $T=2\pi$). Koch and Hershey (10) report that the pool size, N, is "about" 34 in broth media. Burst sizes, $\overline{B}$, are typically about 200 under similar conditions. Neither of these parameters is known precisely and the burst size, at least, varies considerably with different bacterial strains and with different growth conditions. We have assumed the reasonable and convenient value of $T=2\pi$ for the ratio of burst size to pool size. In Table 1, we give the values of $\overline{P}_n/(1 - \overline{P}_0)$ (for n=1 - 20) calculated on the assumption that $T=2\pi$.

In Figure 1 we compare the calculations of Table 1 with the experimental data and theoretical curve of Luria (1). The agreement is quite reasonable, particularly when it is considered that T was not used as a free parameter to obtain a "best fit" but rather was chosen, if somewhat arbitrarily, on other grounds. It is not surprising that the calculated values fall below the data for large clones, for we have neglected the contribution due to mutations which occur during the eclipse period. This too could be calculated, but only at the expense of adding another parameter. Our main purpose is not to attempt to fit the experimental data precisely; it is rather to show that the steady-state pool concept offers no difficulty

in understanding the mutant frequency results.

In Figure 1 we also present the data of Stahl and Streisinger giving the clone size distribution for recombinants between closely linked markers. We wish to thank these workers for their kind permission to use their unpublished data. These data fall below the calculated distribution, even when T is assumed to be π, which may be considered to be a lower limit. There is now good evidence (11) that the reason for the disagreement is that recombinant formation is not a simple instantaneous event. It seems likely that the formation of a recombinant for closely linked markers involves first the formation of a non-recombinant heterozygote, which is then followed by the formation of a recombinant heterozygote, which in turn may segregate pure recombinants after some delay. The net result is that pure recombinants tend to be formed late in the latent period, and accordingly have little chance to multiply. The problem of the formation of recombinants among closely linked markers will be considered in more detail in Part Two of this Thesis.

An old and repeatedly confirmed observation is that the variance of the phage yields from single cells is large (12). This large variance in burst size is undoubtedly due to many causes. For instance, under most conditions of growth, the host bacterial cells are of widely varying sizes. Perhaps, other properties of the cells vary

similarly. It should be pointed out, however, that the model for phage replication that we have used here leads to a quite large variance in burst size, even if other sources of variability were absent. We are indebted to Dr. R. H. Epstein for providing the raw data from his "zero-dose" single burst experiment reported in (13). In the experiments of Epstein, special conditions were used which eliminate at least part of the usual cellular variability. When these data are corrected for multiple bursts, we obtain a value of 151 for the mean and a value of 7,986 for the variance. This gives a ratio of variance to mean of 19. From equation 24 (and again assuming a value for $T=2\pi$), we calculate a ratio of about 21. Thus we conclude that a steady-state model in which phage growth and maturation are essentially random processes is sufficient to account for the observed variance in burst size.

TABLE 1

The distribution of clone sizes of mutants originating in a steady-state pool of vegetative phage. The values for n = 0 - 2 were calculated from the appropriate derivatives of G(s,T) (eq. 11,13). The values for n = 4 - 20 were calculated from the series approximation (eq. 16). The value for n = 3 was calculated by both methods; these calculated values differed by less than 0.1%. The parameter T is assumed to be 2π.

Clone size	Frequency	Clone size	Frequency
n	$\bar{P}_n / (1 - \bar{P}_0)$	n	$\bar{P}_n / (1 - \bar{P}_0)$
1	0.562	11	0.00847
2	0.140	12	0.00727
3	0.0702	13	0.00630
4	0.0438	14	0.00549
5	0.0306	15	0.00482
6	0.0229	16	0.00425
7	0.0179	17	0.00376
8	0.0144	18	0.00334
9	0.0119	19	0.00298
10	0.00996	20	0.00266

FIGURE 1

Luria plots of experimental and theoretical clone size distributions. The abscissa is the clone size, n; the ordinate is the frequency with which clones of size n or greater are found. Both scales are logarithmic. Filled circles are the unpublished data of Stahl and Streisinger (see text) on recombinants for very closely linked markers. Open circles are the data of Luria (1) on spontaneous mutants. The broken line is the theoretical distribution of Luria. The three unbroken curves are calculated from the steady-state pool model presented here; each curve is labeled with the value of the parameter T (ratio of burst size to pool size) which was used for calculation.

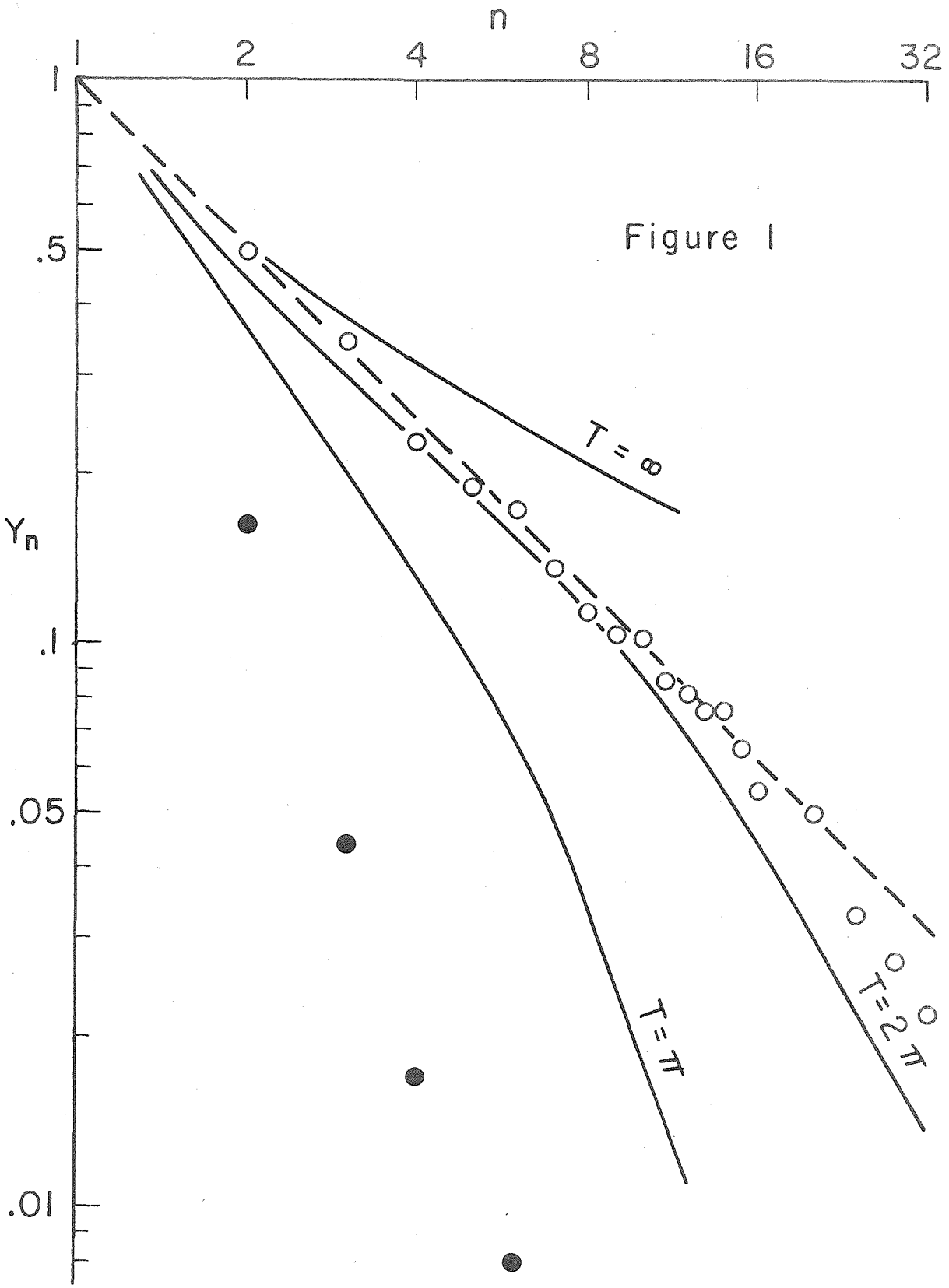


Figure 1

II. RANDOM MATING THEORY

1. Introduction.

The essential content of this chapter is embodied in an article, "The Theory of Formal Phage Genetics" by Charley Steinberg and Frank Stahl, published in the Cold Spring Harbor Symposium for 1958 (pp. 42-45). This article was appended to a review of phage genetics by Hershey (4) and, in consequence, contains little background material. Hence the need for the present introductory remarks.

When a bacterium is mixedly infected with phages which are closely related but which differ by two or more genetic markers, it is observed that genetically recombinant particles occur in the progeny. This phenomenon was first reported by Delbrück and Bailey in 1946 (14), and many of the basic features were already established by 1949 from the work of Hershey and Rotman (15). A standard cross consists of mixedly infecting a population of bacteria with high, equal multiplicities of two genetically marked parental phage strains. After infection, growth is allowed to proceed until normal lysis occurs, and the frequencies of the various genotypes among the progeny are determined. When the parents differ by only two genetic factors, the two recombinant types occur in equal frequency*

---

*This is a trivial fact. If the two recombinant types were not equal in frequency, then one of the markers would have been at a "selective disadvantage" and would not have been used for such studies.

in the progeny, so that there is only one parameter, usually chosen to be the total frequency of recombinants. As with higher organisms, such recombination frequencies can be taken to represent a measure of a distance, and linear genetic maps can be constructed. In fact, provided certain technical requirements are fulfilled (16), the frequencies are sufficiently reproducible to permit the construction of maps from two-factor crosses alone. Three-factor crosses, however, provide more definitive information for the ordering of markers. Thus the underlying genetic structure of phages appears to be quite similar to that of higher organisms.

Considering for the moment only standard crosses, there are two facts, however, which would seem to require explanation. (a) Saturation. For the related phages T2 and T4, the maximum frequency of recombination in two-factor crosses is about 40% (16). This is significantly different from either the 50% usually found in higher organisms or the 25% which would be predicted for a homothallic organism. Furthermore, this maximum frequency differs for different phages, being about 15% for phages T1 (17) and lambda (18). (b) Negative Interference. In three-factor crosses, the two possible recombinations are positively correlated; that is, the exchanges seem to be "clustered." This phenomenon is usually called "negative interference." It is not surprising, perhaps, that the

degree of negative interference is greatest with those phages which show the lowest maximum recombination frequencies.

Still more striking features of the recombination process are revealed when other than standard crosses are performed. In particular, more than two parental genotypes may be used, or the parents may be used in very unequal proportions, or the time of lysis may be artificially decreased or increased. (a) Triparental Recombinants. When three distinctively marked parents are used in a cross, the progeny contain particles which have derived markers from each of the three parents (15). The frequency of such triparental recombinants is not low; in fact, the two exchanges are uncorrelated (19). (b) Loss of Minority Parent. When the multiplicity of infection of the two parents (in an otherwise standard cross) are grossly unequal, the progeny may contain more recombinant particles than particles with the genotype of the minority parent (20). (c) Genetic Drift. When the infected cells are lysed prematurely, the frequency of recombinants is found to increase with time of lysis (20). This is true even for freely recombining markers (i.e., markers which show the maximum frequency of recombination). If lysis is inhibited beyond the normal time, the frequency of recombinants increases still further. For rather closely linked markers, this increase is linear in time (21).

The experimental facts summarized above certainly demonstrate that the phage cross is not a straightforward analogue of a simple genetic cross as it is known in higher organisms. In fact, either the finding of triparental recombinants or the phenomenon of the loss of the minority parent would alone be sufficient to dispel such a notion. Since the time of the Ptolemaic system, at least, it has been human nature, when faced with the inadequacy of a simple model, to inquire whether an iteration of the model might not be adequate. Visconti and Delbrück (7) in 1953 suggested that in the vegetative pool (see Chapter I, section 1) there occur repeated, pairwise matings between genetically complete vegetative particles. They further assumed that such matings were at random with respect to partners, and that recombinants were formed by a mechanism which may be described as two-strand crossing-over. Even at a qualitative level, the Visconti-Delbrück theory accounts in an elegant and pleasing fashion for all of the experimental observations we have described. These authors formulated a detailed mathematical treatment of the consequences of their model for a three-factor cross which gives the frequencies of genotypes in the progeny as a function of their frequency in the parental population, the recombination probabilities for a single hypothetical mating event, and the average number of matings which a progeny particle has experienced ("rounds of mating").

They also executed a series of experiments designed to estimate the parameters and to test further the validity of the theory. Fairly good quantitative agreement between theory and experiments was obtained by assuming an average of about five rounds of mating at normal lysis for the bacteriophage T2. Studies by other workers with phage lambda (18) and phage T1 (17) are compatible with somewhat less than one round of mating.

Our work was motivated first of all by pedagogical considerations. We wanted to formulate the Visconti-Delbrück theory by an approach which dealt directly with probabilities without the intervention of recursion formulae or differential equations. Secondly, there was the question of the mechanism of recombination, which Visconti and Delbrück had assumed to be two-strand crossing-over. But the model which was popular at the time of our work was the so-called copy-choice mechanism, the most fashionable version being that of Levinthal (22). Some workers apparently believed intuitively that the same formal structure would result whatever the mechanism (23), while others believed quite the opposite (24). Thus we find Stent (25) using the fact that there are five rounds of mating as a principal argument for a theory of growth and recombination which assumes quite another mechanism of recombination. One of the principal results of our investigations is that the mechanism of recombination influences

the theory only to the extent of a temporal scale factor. Hence, the formal structure of the theory does not depend upon the mechanism of recombination, but the interpretation of the numerical value of the average rounds of mating does depend upon this mechanism.

At this point, Dr. A. D. Hershey communicated to us some new results on the stimulation of recombination in phage T2 by ultraviolet light (4). Hershey felt that his results were incompatible with the notion that recombination occurred *via* pairwise matings but were compatible with a "group-mating" process. The same idea was put forth independently by Bresch and Starlinger (26). In hindsight it is obvious that our picture of a vegetative phage is much more analogous to a chromosome than to an independent organism, and *ipso facto* an infected cell is analogous to a highly polyploid nucleus. We recall that while synapsis of chromosomes in a triploid organism is always pairwise at any given level, the synaptic figure as a whole is usually trivalent. And we further recall that Bridges and Anderson (27) recovered, in the progeny from triploid flies, chromosomes which had genetic markers derived from all three chromosomes of a homologous set. In any event, using the mathematical approach developed for the Visconti-Delbrück model, it was quite simple to extend the theory to cover group-mating. We are now left with

essentially two non-trivial assumptions: that in some sense it is possible to break the process of recombinant production into a series of independent events or matings, and that the participants in a mating are chosen at random from the entire vegetative pool.

2. Theory.

The Theory of Formal Phage Genetics

Charley Steinberg and Frank Stahl[2]

Division of Biology, California Institute of Technology, Pasadena, California

This appendix will consist of three sections. In the first section, we shall look at the older, pairwise (Visconti-Delbrück) mating theory in a somewhat new way, which deals directly with probabilities without the intervention of recursion formulae or differential equations. Because of this we feel that the physical picture of phage recombination is more clearly embodied in the formalism, and that the theory is more easily generalizable. In the second section we present the formalism for a new theory of recombination—group mating theory—the idea for which was first brought to our attention by Dr. A. D. Hershey. In the third section we shall present a more general theory which encompasses the first two as special cases.

Pair Mating Theory

This theory was originally proposed by Visconti and Delbrück (1953). As we shall see, however, their presentation was unnecessarily restrictive, and the same formal theory can describe models proposed by several other authors.

We picture a sequence of successive matings between *pairs* of hypothetical entities which we call vegetative phage. But first, let us look at a single mating.

We suppose that each mating occurs in a "black box"; *i.e.*, we do not try to draw a detailed picture of the recombination mechanism within a single mating. It is precisely here that we gain generality since our results will be valid whatever the detailed mechanism of the elementary recombination process may be.

The phage which enter into a mating may differ with respect to allelic state at one or more loci. However, we shall make the conventional assumption that the events which go on in the "box" between two phages are not influenced by the markers. In particular we mean that it does make sense to talk about recombination between loci, even when the mating phages are isogenic. For this reason, we forget the alleles for the moment, and prior to the initiation of *every* mating act we mark each of the three loci (1, 2, and 3) of one parent with red paint; we use yellow paint to mark each of the three loci of the other parent. N phages emerge from each mating; each of the N phages has a mark either red or yellow at each locus depending on whether the gene at that locus was derived from the red or yellow parent respectively.

[2] Present address: *Department of Zoology, University of Missouri, Columbia, Missouri.*

With three loci each of the N phages is identical in color to one of the parents, or with some probability differs from one parent at only one locus, *i.e.*, *looks as if* it has been "color-converted" at just one locus.

Locus	1	2	3	1	2	3
Parents	R	R	R	Y	Y	Y
Types of phage	R	R	R	Y	Y	Y
which may	Y	R	R	R	Y	Y
emerge from a	R	Y	R	Y	R	Y
mating.	R	R	Y	Y	Y	R

We say that a phage emerging from a mating is *descended* from the parent it more nearly resembles with respect to color. We are thus able for any phage to trace back in time along a non-branching line of descent originated by a unique infecting particle. We call this line of descent a *lineage*.

We now seek to write that probability that a phage has a specified genotype in terms of the *color-conversion history* of its lineage. For any phage pulled randomly from the population we examine the color-conversion history of its lineage and assign that phage to one of the eight categories tabulated below. For each category we indicate whether a lineage in that category has had no color-conversion at the locus by a "0", or whether it has had one *or more* at that locus by a "1". We denote the probability of a given category by P with appropriate subscripts.

Category	Color-conversion at locus			Probability of category
	1	2	3	
I	0	0	0	P_{000}
II	1	0	0	P_{100}
III	0	1	0	P_{010}
IV	0	0	1	P_{001}
V	0	1	1	P_{011}
VI	1	0	1	P_{101}
VII	1	1	0	P_{110}
VIII	1	1	1	P_{111}

We let the frequencies at time zero of any given genotype ijk be designated by a_{ijk}. We designate by b_{ijk} the frequency of the given genotype in the population at some later time. If we wish to specify the frequency of a partial genotype we write a "dot" at those loci whose allelotype is immaterial; *e.g.*, $a_{ij.}$ is the frequency at time zero of phage with the specified alleles i and j at the first two loci, but which are unspecified as to allelotype at the third locus.

We now consider the contribution to b_{ijk} of each category in the above table. This contribution will be the product of the following probabilities:

(1) The probability that for each locus color-converted one or more times in the lineage, the *last* conversion was in a mating with a phage carrying the specified allele at the converted locus. For each locus, this probability is just the frequency in the pool of the specified allele.

(2) The probability that the lineage was initiated by an input phage carrying the specified alleles at *each* of the unconverted loci. This probability is just the frequency at time zero of the appropriate phage.

(3) The product of these probabilities is then weighted by the probability of the category.

Category	Contribution to b_{ijk}
I	$a_{ijk}P_{000}$
II	$a_{i\cdot\cdot}a_{\cdot jk}P_{100}$
III	$a_{\cdot j\cdot}a_{i\cdot k}P_{010}$
IV	$a_{\cdot\cdot k}a_{ij\cdot}P_{001}$
V	$a_{\cdot j\cdot}a_{\cdot\cdot k}a_{i\cdot\cdot}P_{011}$
VI	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{101}$
VII	$a_{i\cdot\cdot}a_{\cdot j\cdot}a_{\cdot\cdot k}P_{110}$
VIII	$a_{i\cdot\cdot}a_{\cdot j\cdot}a_{\cdot\cdot k}P_{111}$

The desired probability b_{ijk} is the sum of these eight terms. This is the most general expression for a three locus problem. We have put no restriction on the distribution of the number of matings among lineages, or on the mechanism of recombination (conversion, crossing over, copy choice all with or without interference, positive or negative). We have put no restriction on the number of alleles at each locus.

We now specify the P's in terms of the "mechanism of recombination," the color-conversion probabilities, and the number of matings per lineage.

A fraction F of the N phages emerging from a mating have a probability greater than zero of being one of the converted types. Various combinations of F and N correspond to hypothetical mechanisms of recombination. For instance:

	Break-reunion		Copy-choice	
	2 make 1	2 make 2	2 make 3	2 make 4
Reference	Levinthal 1953	Visconti and Delbrück 1953	Levinthal 1953	Hershey, this Symposium
N	1	2	3	4
F	1	1	$\frac{1}{3}$	$\frac{1}{2}$

To each of the FN phage (which can be convertant) emerging from a mating we assign a probability that it has been converted in that mating at a given locus.

Locus converted	Probability
None	c_0
1	c_1
2	c_2
3	c_3

It follows from the definition of F and the c_n's that the *average* probability that a given phage issuing from a mating has been converted at the nth locus in that mating is Fc_n.

We may now specify the P's in terms of m (matings per lineage), F, and the c_n's according to the distribution of mating acts among lineages. For instance, for a Poisson distribution (random-in-time mating; Visconti and Delbrück, 1953) we have P values which are the products of zeroth and 1 − zeroth terms of the Poisson expression; *e.g.*,

$$P_{110} = (1 - e^{-mFc_1})(1 - e^{-mFc_2})\, e^{-mFc_3}$$

$$P_{001} = e^{-mFc_1}\, e^{-mFc_2}\, (1 - e^{-mFc_3})$$

etc.

Note that F occurs only in the product mF, and that therefore (for random-in-time mating) there is no way of distinguishing (by three factor crosses) between the possibilities of 2 make 2, 2 make 3, etc. The above expressions for the P's may be corrected for spread in maturation according to equation 11 of Visconti and Delbrück (1953).

The c_n's may be specified in terms of recombination probabilities for a linear structure as follows:

$$c_1 = p_{12}(1 - Sp_{23})$$

$$c_2 = Sp_{12}p_{23}$$

$$c_3 = p_{23}(1 - Sp_{12})$$

where

p_{12} is the probability of recombination between the loci 1 and 2

p_{23} is the probability of recombination between the loci 2 and 3

S is the coefficient of coincidence.

Group Mating Theory

The kinetics for group mating can be written similarly to that for pairwise matings. We picture a group mating in the following way. A large number of phages constituting a random sample of the population enter each act. As before, we color the three loci on each of the parents *distinctively*. N phages emerge from the act. Each of the N phages has a color mark at each locus corresponding to the color of the parent from which the gene at that locus was derived.

Locus	1 2 3	1 2 3	1 2 3	
Parents	*R R R*	*Y Y Y*	*G G G*	*etc.*
Types of phage which may emerge from a group interaction.	*R R R*	*Y Y Y*	*G G G*	*etc.*
	· R R	*· Y Y*	*· G G*	*etc.*
	· R ·	*· Y ·*	*· G ·*	*etc.*
	R R ·	*Y Y ·*	*G G ·*	*etc.*

We say that a phage emerging from a group mating is descended from the parent which it resembles with respect to color at locus 2. A lineage is a line of descent. *Each* group mating can have four outcomes for a given lineage: (1) no conversion in that act; (2) conversion at locus 1 only; (3) conversion at loci 1 and 3; (4) conversion at locus 3 only. We now seek to write the probability that a phage has a specified genotype in terms of the color-conversion history of its lineage.

For any phage pulled randomly from the population, its lineage must correspond to one of the eight categories tabulated below. For each category we indicate whether a lineage in that category has had no color conversions at the indicated locus (loci) by a "0", or whether it has had one *or more* at that locus (loci) by a "1". We denote the probability of a given category by P with appropriate subscripts.

Category	Conversion at locus (loci)			Probability of category
	1 only	1 & 3	3 only	
I	0	0	0	P_{000}
II	1	0	0	P_{100}
III	0	1	0	P_{010}
IV	0	0	1	P_{001}
V	0	1	1	P_{011}
VI	1	0	1	P_{101}
VII	1	1	0	P_{110}
VIII	1	1	1	P_{111}

We now consider the contribution of each category in the table to the desired frequency b_{ijk}. For the general case, this contribution will be the product of the following probabilities:

(1) The probability that for each locus color-converted one or more times in the lineage, the last conversion was a result of an interaction with a phage carrying the specified allele at the converted locus. For each locus, this probability is just the frequency in the pool of the specified allele.

(2) The probability that the lineage was initiated by an input phage carrying the specified alleles at each of the unconverted loci (remembering that locus 2 cannot convert, by definition of descent).

(3) The product of these probabilities is then weighted by the probability of the category.

Category	Contribution to b_{ijk}
I	$a_{ijk}P_{000}$
II	$a_{i\cdot\cdot}a_{\cdot jk}P_{100}$
III	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{010}$
IV	$a_{\cdot\cdot k}a_{ij\cdot}P_{001}$
V	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{011}$
VI	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{101}$
VII	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{110}$
VIII	$a_{i\cdot\cdot}a_{\cdot\cdot k}a_{\cdot j\cdot}P_{111}$

The desired probability b_{ijk} is the sum of these eight terms.

We now specify the P's in terms of the "mechanism of recombination," the color-conversion probabilities, and the number of group matings per lineage.

A fraction F of the N phage emerging from a group mating have a probability greater than zero of being one of the converted types. We distinguish between "break-reunion" and "copy-choice" for which $F = 1$ and $F < 1$, respectively.

To each of the FN phage (which can be convertant) emerging from a group mating, we assign a probability that it has been converted in the mating at a given locus (loci).

Locus (loci) converted	Probability
none	c_0
1	c_1
1 and 3	c_2
3	c_3

Specification of the P's in terms of m, F, and the c_n's, and the specification of the c_n's in terms of a linear linkage structure are formally identical to those of the pair mating theory. A consideration of the applications of these expressions to the case of freely recombining markers will serve to illustrate the difference in the meaning of the parameters in the two theories.

For both pair mating and group mating m, F, and c_n occur always as a product. The disentanglement of mF from c is possible whenever *a priori* knowledge about the c's is available. If we find three loci which behave symmetrically with respect to each other, we may presume that they are freely recombining. For the pair mating case, this implies that $c_0 = c_1 = c_2 = c_3 = \frac{1}{4}$. For group mating this implies $c_2 = 1$, all other c_n's = zero.

Note that p_{12} and p_{23} are defined differently for pair mating and group mating. For pair mating, p varies from 0 to 0.5; for group mating p ranges from 0 to 1.

Generalized Theory

Here we wish to write a more general theory which includes pair and group mating theories as special cases. We assume as before that a group of vegetative phages participate in a mating, but relax the requirements that all participate equally and that the group must be large enough to include a statistically representative sample of the entire population.

The presentation will be schematic, using the concepts and notation of the preceding treatments.

Segregation Types

	Example	Definition	Probability
0	*RRR*	all three markers from one parent	Fc_0
1	*YRR*	markers 2 & 3 from one parent, marker 1 from another	Fc_1
2	*RYR*	markers 1 & 3 from one parent, marker 2 from another	Fc_2
3	*RRY*	markers 1 & 2 from one parent, marker 3 from another	Fc_3
4	*YRB*	each marker from a different parent	Fc_4

Definition of Descent

If the offspring derived two or three markers from one parent, it is descended from that parent; if the three markers are all derived from different parents, the offspring is descended from the parent which donated the middle marker. (In the example above, all are descended from the *RRR* parent.)

Classification of Lineages

We classify lineages according to whether there are none or more than none segregations of each of the types 1, 2, 3, and 4. There are thus 16 categories of lineages; the probability of each category is denoted by P with 4 subscripts in a manner analogous to our preceding treatments.

Example: P_{1001} is the probability that a phage particle has a lineage belonging to the category characterized by one or more type 1 segregations, no type 2 segregations, no type 3 segregations, *and* one or more type 4 segregations.

The contribution of each category to the genotype frequency b_{ijk} is the probability of the category multiplied by the appropriate coefficient, according to the table:

Category	Coefficient	Probability of category	Segregations of type: 1	2	3	4
I	a_{ijk}	P_{0000}	0	0	0	0
II	$a_{i\cdot\cdot}a_{\cdot jk}$	P_{1000}	1	0	0	0
III	$a_{\cdot j\cdot}a_{i\cdot k}$	P_{0100}	0	1	0	0
IV	$a_{\cdot\cdot k}a_{ij\cdot}$	P_{0010}	0	0	1	0
V	A	P_{0110}	0	1	1	0
VI	A	P_{1010}	1	0	1	0
VII	A	P_{1100}	1	1	0	0
VIII	A	P_{1110}	1	1	1	0
IX	A	P_{0001}	0	0	0	1
X	A	P_{1001}	1	0	0	1
XI	A	P_{0101}	0	1	0	1
XII	A	P_{0011}	0	0	1	1
XIII	A	P_{0111}	0	1	1	1
XIV	A	P_{1011}	1	0	1	1
XV	A	P_{1101}	1	1	0	1
XVI	A	P_{1111}	1	1	1	1

Notes: $A = a_{i\cdot\cdot}a_{\cdot j\cdot}a_{\cdot\cdot k}$
"1" means one *or more*.

The specification of the P's in terms of m, F, and the c_n's are again analogous to those given in the pairwise theory. Note however that the sum of the contributions to b_{ijk} from categories IX–XVI is $A(1 - e^{-mFc_4})$ and that the sum of the contributions from categories I–XIII is simply the solution for pairwise mating multiplied by e^{-mFc_4}.

The specification of the c's depends upon the model. For recombination between linear structures:

$$c_1 = p_1(1 - p_2S)$$
$$c_2 = \alpha p_1p_2S$$
$$c_3 = p_2(1 - p_1S)$$
$$c_4 = \beta p_1p_2S$$

and $\alpha + \beta = 1$.

If exactly M phages participate *equally* in a mating:

$$\alpha = \frac{1}{M - 1} \qquad \beta = \frac{M - 2}{M - 1}$$

For $M = 2$, $c_4 = 0$, and we have the pairwise interaction model; for $M \to \infty$, $c_2 \to 0$, and we have the group interaction model.

We feel it preferable to leave α an arbitrary parameter since it would be influenced by factors which are impossible to estimate at this time, *e.g.*, a distribution in M, or an unequal participation of phages in a group mating perhaps due to geometrical considerations.

We have seen that for freely recombining markers, $p = \frac{1}{2}$ and $p = 1$ for pair and group mating respectively. For the general case it is reasonable that p for freely recombining markers should fall between $\frac{1}{2}$ and 1; in fact $p = 1/(\alpha + 1)$ for such markers.

References

Barricelli, N. A., 1956, A "chromosomic" recombination theory for multiplicity reactivation in phages. Acta Biotheoretica *11:* 107–120.

Baylor, M. B., Hurst, D., Allen, S., and Bertani, E., 1957, The frequency and distribution of loci affecting host range in the coliphage T2H. Genetics *42:* 104–120.

Benzer, S., 1955, Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci. Wash. *41:* 344–354.

1957, The elementary units of heredity. In: The Chemical Basis of Heredity, W. D. McElroy and B. Glass, eds. Johns Hopkins Press, Baltimore.

Bresch, C., 1955, Zum Paarungsmechanismus von Bakteriophagen. Z. Naturforsch. *10b:* 545–561.

Bresch, C., and Trautner, T., 1955, Zur Kinetic der Recombinantbildung bei T1-Bacteriophagen. Z. Naturforsch. *10b:* 436–440.

Chase, M., and Doermann, A. H., 1958, High negative interference over short segments of the genetic structure of bacteriophage T4. Genetics, (in press).

Delbrück, M., 1958, Bacteriophage genetics. Proc. 4th Intern. Poliomyelitis Conf., (in press).

Delbrück, M., and Stent, G. S., 1957, On the mechanism of DNA replication. In: The Chemical Basis of Heredity, W. D. McElroy and B. Glass, eds. Johns Hopkins Press, Baltimore.

Doermann, A. H., 1953, The vegetative state in the life cycle of bacteriophage: evidence for its occurrence and its genetic characterization. Cold Spring Harbor Symposia Quant. Biol. *18:* 3–11.

Doermann, A. H., and Hill, M., 1953, Genetic structure of bacteriophage T4 as described by recombination studies of factors influencing plaque morphology. Genetics *38:* 79–90.

Doermann, A. H., Chase, M., and Stahl, F. W., 1955, Genetic recombination and replication in bacteriophage. J. Cellular Comp. Physiol. *45:* Supp. 2, 51–74.

Dulbecco, R., 1952, A critical test of the recombination theory of multiplicity reactivation. J. Bacteriol. *63:* 199–207.

Edgar, R. S., 1956. See discussion of paper by Streisinger and Franklin (1956).

1958a, Phenotypic properties of heterozygotes in the bacteriophage T4. Genetics, (in press).

1958b, personal communication.

Edgar, R. S., and Steinberg, C. M., 1958, On the origin of high negative interference over short segments of the genetic structure of bacteriophage T4. Genetics, (in press).

EPSTEIN, R. H., 1957, A study of multiplicity reactivation in the bacteriophage T4. Dissertation, University of Rochester.

FLAKS, J. G., and COHEN, S., 1957, The enzymic synthesis of 5-hydroxymethyldeoxycytidylic acid. Biochim. et Biophys. Acta *25:* 667–668.

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

HARM, W., 1956, On the mechanism of multiplicity reactivation. Virology *2:* 559–564.

HERSHEY, A. D., 1952, Reproduction of bacteriophage. Intern. Rev. Cytol. *1*, 119–134.

1953a, Inheritance in bacteriophage. Advances in Genet. *5:* 89–106.

1953b, Nucleic acid economy in bacteria infected with bacteriophage T2. II. Phage precursor nucleic acid. J. Gen. Physiol. *37:* 1–23.

1957, Some minor components of bacteriophage T2 particles. Virology *4:* 237–264.

HERSHEY, A. D., and BURGI, E., 1956, Genetic significance of the transfer of nucleic acid from parental to offspring phage. Cold Spring Harbor Symposia Quant. Biol. *21:* 91–101.

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

HERSHEY, A. D., and Melechen, N. E., 1957, Synthesis of phage precursor nucleic acid in the presence of chloramphenicol. Virology *3:* 207–236.

HERSHEY, A. D., and ROTMAN, R., 1949, Genetic recombination between host range and plaque type mutants of bacteriophage in single bacterial cells. Genetics *34:* 44–71.

HERSHEY, A. D., BURGI, E., and STREISINGER, G., 1958, Genetic recombination between phages in the presence of chloramphenicol. Virology, (in press.)

JACOB, F., and WOLLMAN, E., 1954, Etude génétique d'un bactériophage tempéré d'*Escherichia coli*. I. Le système génétique du bactériophage lambda. Ann. inst. Pasteur *87:* 653–674.

1955, Etude génétique d'un bactériophage tempéré d'*Escherichia coli*. III. Effet du rayonnement ultraviolet sur la recombinaison génétique. Ann. inst. Pasteur *88:* 724–749.

1957, Genetic aspects of lysogeny. In: The Chemical Basis of Heredity, W. D. McElroy and B. Glass, eds. Johns Hopkins Press, Baltimore.

KAISER, A. D., 1955, A genetic study of the temperate coliphage lambda. Virology *1:* 424–443.

KRIEG, D. R., 1957, A study of gene action in ultraviolet-irradiated bacteriophage. Dissertation, University of Rochester.

LEDERBERG, J., 1955, Recombination mechanisms in bacteria. J. Cellular Comp. Physiol. *45:* Suppl. 2, 75–107.

LENNOX, E. S., LEVINTHAL, C., and SMITH, F., 1953, The effect of finite input in reducing recombinant frequency. Genetics *38:* 508–511.

LEVINTHAL, C., 1954, Recombination in phage T2: its relationship to heterozygosis and growth. Genetics *39:* 169–184.

1956, The mechanism of DNA replication and genetic recombination in phage. Proc. Natl. Acad. Sci. Wash. *42:* 394–404.

LEVINTHAL, C., and VISCONTI, N., 1953, Growth and recombination in bacterial viruses. Genetics *38:* 500–511.

LURIA, S. E., 1947, Reactivation of irradiated bacteriophage by transfer of self-reproducing units. Proc. Natl. Acad. Sci. Wash. *9:* 253–264.

1951, The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage production. Cold Spring Harbor Symposia Quant. Biol. *16:* 463–470.

LURIA, S. E., and DULBECCO, R., 1949, Genetic recombinations leading to production of active bacteriophage from ultraviolet inactivated bacteriophage particles. Genetics *34:* 93–122.

LURIA, S. E., and LATARJET, R., 1947, Ultraviolet irradiation of bacteriophage during intracellular growth. J. Bacteriol. *53:* 149–163.

MCFALL, E., and STENT, G. S., 1958, Three star mutants of coliphage T2. J. Gen. Microbiol. *18:* 346–363.

MESELSON, M., and STAHL, F. W., 1958, The replication of DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. Wash. *44*, (in press).

MESELSON, M., STAHL, F. W., and VINOGRAD, J., 1957, Equilibrium sedimentation of macromolecules in density gradients. Proc. Natl. Acad. Sci. Wash. *43:* 581–583.

STENT, G. S., 1958, Mating in the reproduction of bacterial viruses. Advances in Virus Research *5:* 95–149.

STENT, G. S., JERNE, N., and SATO, G., 1957. See Delbrück and Stent, 1957.

STREISINGER, G., and FRANKLIN, N., 1956, Mutation and recombination at the host range genetic region of phage T2. Cold Spring Harbor Symposia Quant. Biol. *21:* 103–109.

SYMONDS, N., 1957, Effects of ultraviolet light during the second half of the latent period on bacteria infected with phage T2. Virology *3:* 485–495.

TESSMAN, I., TESSMAN, E. S., and STENT, G. S., 1957, The relative radiosensitivity of bacteriophages S13 and T2. Virology *4:* 209–215.

TOMIZAWA, J., 1958, Sensitivity of phage precursor nucleic acid, synthesized in the presence of chloramphenicol, to ultraviolet irradiation. Virology, (in press).

VISCONTI, N., and DELBRÜCK, M., 1953, The mechanism of genetic recombination in phage. Genetics *38*, 5–33.

WATSON, J. D., and CRICK, F. H. C., 1953, A structure for deoxyribonucleic acid. Nature, Lond. *171:* 737–738.

WILKINS, M. F. H., STOKES, A. R., and WILSON, H. R., 1953, Molecular structure of deoxypentose nucleic acids. Nature, Lond. *171:* 738–740.

WOLLMAN, E. L., and JACOB, F., 1954, Etude génétique d'un bactériophage tempéré d'*Escherichia coli*. II. Mécanisme de la recombinaison génétique. Ann. inst. Pasteur. *87:* 674–691.

ZAHLER, S. A., 1958, Some biological properties of bacteriophages S13 and ϕ X-174. J. Bacteriol. *75:* 310–315.

3. Discussion.

The principal results of the theory bear upon what we have termed the mechanism of recombination and the size of the mating group. The mechanism of recombination influences the formal equations only to the extent of a scale factor, thus precluding the possibility that the usual types of mass-lysate experiments could yield critical information concerning this question. Concerning the size of the mating group, the situation is different. Within the theoretical framework here presented, it is possible to distinguish experimentally between the pair- and group-mating alternatives. Toward this end, Hausmann and Bresch (28) performed an elegant series of experiments with phage T1, but obtained inconclusive results. Hershey (4) concluded that his experiments on the stimulation of recombination in T2 and T4 by ultraviolet light could be most easily explained on the basis of group-mating; but, of course, any interpretation depends upon our understanding of the biological effects of ultraviolet light. Millard Susman (29), of this Institute, has obtained critical results with phage T4 which are in substantial agreement with the notion of pair-mating. Thus the experimental picture is somewhat confusing.

But behind the experimental confusion lies an even more serious theoretical confusion. Many workers, the most eloquent being Bresch and his collaborators (30),

have questioned the assumption that the participants in a mating are chosen at random from the entire vegetative pool. More specifically, it is suggested that the products of a mating are likely to remain spacially contiguous and are therefore more likely to mate among themselves than with other particles in the pool. This notion is referred to as "topography" or "poor mixing."

It is easy to show that the question of topography is confounded with the question of the size of the mating group. Let us consider the maximum amount of information that can be extracted from the usual types of mass-lysate recombination experiments which the theory was written to describe. For a three-factor cross, it is in principle possible to determine the fraction of the progeny in which all three markers are descended from one parental particle, the fraction in which the markers are descended from three different parents, and the three fractions in which one of the markers is descended from one parent and the other two markers are descended from another parental particle. Furthermore, these frequencies may all be determined as a function of the time of lysis. An experiment will usually yield less information than the above, but it cannot yield more. For the pair- vs. group-mating alternative, the relevant subclass of the progeny is the one in which the central marker alone is descended from one parental particle. Among this subclass, we determine the relative frequencies

with which the outside markers are descended from one parental particle and from two parental particles. These relative frequencies are then compared with some theoretical expectation. But if it is found that a relatively large proportion have outside markers descended from a single parental particle, is it because mating is intrinsically pairwise, or is it because the locale in which a given mating event takes place is unlikely to contain the descendents of more than two parental particles?

The notion of topography is quite complex, and no one has succeeded in any reasonably complete formulation of its detailed consequences. With the T-even phages, at least, it is possible to use presently available data to preclude the possibility of certain crude but specific models. It does not seem worthwhile to carry the discussion further, however, for the questions of topography and the size of the mating group are still very much open.

REFERENCES

1. Luria, S. E., 1951. The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. Cold Spring Harbor Symposia on Quantitative Biology, XVI: 463-478.

2. Luria, S. E. and M. Delbrück, 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28: 491-511.

3. Armitage, P., 1953. Statistical concepts in the theory of bacterial mutation. Journal of Hygiene, 51: 162-184.

4. Hershey, A. D., 1958. The production of recombinants in phage crosses. Cold Spring Harbor Symposia on Quantitative Biology, XXIII: 19-46.

5. Stahl, F. W., 1956. The effects of the decay of incorporated radioactive phosphorus on the genome of bacteriophage T4. Virology, 2: 206-234.

6. Levinthal, C., 1956. Personal communication.

7. Visconti, N., and M. Delbrück, 1953. The mechanism of genetic recombination in phage. Genetics, 38: 5-33.

8. Feller, W., 1957. An introduction to probability theory and its applications (second edition). Wiley, New York.

9. Bromwich, T. J. I'a., 1926. An introduction to the theory of infinite series. (Second edition, p. 294). Macmillan, London.

10. Koch, G. and A. D. Hershey, 1959. Synthesis of phage-precursor protein in bacteria infected with T2. J. Mol. Biol., 1: 260-276.

11. Edgar, R. S., 1961. High negative interference and heterozygosis: a study of the mechanism of recombination in bacteriophage T4. Virology, 13: 1-12.

12. Delbrück, M., 1945. The burst size distribution in the growth of bacterial viruses. J. Bact., 50: 131-135.

13. Epstein, R. H., 1958. A study of multiplicity--reactivation in bacteriophage T4. I. Genetic and functional analysis of T4D-K12 (λ) complexes. Virology, 6: 382-404.

14. Delbrück, M., and W. T. Bailey, Jr., 1946. Induced mutations in bacterial viruses. Cold Spring Harbor Symposia on Quantitative Biology, XI: 33-37.

15. Hershey, A. D., and R. Rotman, 1949. Genetic recombination between host-range and plaque-type mutants of bacteriophage in single bacterial cells. Genetics, 34: 44-71.

16. Doermann, A. H., and M. Hill, 1953. Genetic structure of bacteriophage T4 as described by recombination studies of factors influencing plaque morphology. Genetics, 38: 79-90.

17. Bresch, C. and T. Trautner, 1955. Zur kinetic der recombinantenbildung bei T1-bacteriophagen. Z. Naturforschung, 10b: 436-440.

18. Kaiser, A. D., 1955. A genetic study of the temperate coliphage lambda. Virology, 1: 424-443.

19. Edgar, R. S., and C. M. Steinberg, 1958. On the origin of high negative interference over short segments of the genetic structure of bacteriophage T4. Virology, 6: 115-128.

20. Doermann, A. H., 1953. The vegetative state in the life cycle of bacteriophage: evidence for its occurrence and its genetic characterization. Cold Spring Harbor Symposia on Quantitative Biology, XVIII: 3-11.

21. Levinthal, C., and N. Visconti, 1953. Growth and recombination in bacterial viruses. Genetics, 38: 500-511.

22. Levinthal, C., 1954. Recombination in phage T2: its relationship to heterozygosis and growth. Genetics, 39: 169-184.

23. Hershey, A. D., 1957. Bacteriophages as genetic and biochemical systems. Advances in Virus Research, 4: 25-61.

24. Levinthal, C., 1958. Personal communication.

25. Stent, G. S., 1958. Mating in the reproduction of bacterial viruses. Advances in Virus Research, 5: 95-149.

26. Bresch, C., and P. Starlinger, 1958. Zum problem der genetischen recombination von bakteriophagen. Die hypothese der kompletten kooperation. Z. f. Vererbungslehre, 89: 459-468.

27. Bridges, C. B., and E. G. Anderson, 1925. Crossing over in the X chromosomes of triploid females of Drosophila melanogaster. Genetics, 10: 418-441.

28. Hausmann, R., and C. Bresch, 1959. Unpublished experiments cited in (30).

29. Susman, M., 1959. Personal communication.

30. Bresch, C., 1959. Recombination in bacteriophage. Annual Rev. Microbiol., 13: 313-334.

PART TWO:

A CRITICAL TEST OF A CURRENT THEORY OF GENETIC RECOMBINATION IN BACTERIOPHAGE

PART TWO

I. INTRODUCTION

1. High Negative Interference.

In an analysis of closely linked markers of bacteriophage T4, Chase and Doermann (1) found a phenomenon which they called "high negative interference" (HNI): the frequency of genetic recombinants whose formation requires multiple exchanges is much greater than would be predicted if exchanges in neighboring regions were statistically independent. This is not a marginal effect. For three-factor biparental crosses in which the outermost markers are less than two map units apart, the fraction of double recombinants exceeds the random expectation by a factor of 14 - 31 . For four-factor crosses the factor is still greater.

Edgar and Steinberg (2) concluded that excess double recombinants due to HNI were produced in a single mating event. This conclusion was based principally upon the results of a series of biparental, three-factor crosses in which the relative frequency of the two parents was systematically varied. Using the type of reasoning common in chemical kinetics, it was demonstrated that the excess recombinants derived their three selected markers from only two parents. It is also possible to design triparental, three-factor crosses so that the selected double

recombinant must derive its three markers from three parents. In this case, there is little, if any, negative interference.*

2. Possible Explanations for HNI.

Since HNI is such a striking phenomenon, we will center our discussion of recombination over short distances around the question: how can we account for HNI? We can conceive of three general directions in which an answer can be sought.

The first possibility might be called The Trivial Hypothesis. By this we mean an explanation which has little to do with the mechanism of recombination per se. As examples, we might mention technical artifacts--that the experiments do not really measure what we think they measure--and various types of statistical inhomogeneities. We have already seen (Part One, Chapter II) that inhomogeneity of "mating experience" can satisfactorily account for the small degree of negative interference found for crosses involving loosely linked markers. It is easy to imagine more insidious types of inhomogeneities which would be important only when relatively rare events are being

---

*Edgar and Steinberg reported that negative interference does occur in triparental crosses, but that the effect is much smaller than in biparental crosses. It was subsequently discovered that these experiments had been improperly executed. Using the same set of markers used for the "unequal input crosses," no significant negative interference was found in triparental crosses (3). These experiments were not extensive enough to conclude that there is no interference at all, but if there is any, it is not HNI.

considered. A particularly simple example would involve a small fraction of infected bacteria which liberate large bursts of recombinants. This can be ruled out by the results of single burst experiments. In fact, as previously pointed out (Part One, Chapter I) the surprising result here is that there are so few large bursts. In the triparental crosses cited above, recombinants are very rare indeed; thus mere rarity itself is not a sufficient condition for HNI.

The second possibility may be entitled The Hypothesis of the Deus ex Machina. In this type of hypothesis, HNI is assumed to be a more or less direct reflection of the mechanism of recombination among closely linked markers, but this mechanism bears little or no relation to the mechanism of recombination among loosely linked markers. Chase and Doermann (1) were able to rule out one model of this type, which they termed "muton conversion": the wild-type "allele" of a marker converts the mutant allele to the wild-type state at a rate which is independent of the other markers in the cross.* Essentially untested, however, is the suggestion of Hershey (5) that exchanges are stimulated by the markers themselves; hence, an excess of multiple

---

*This definition seems to be the content of what is usually termed "gene conversion." Some authors (4), however, insist that the term is simply a name for all of the aberrant phenomena associated with intragenic recombination. Hence our preference for the newer term, muton conversion.

exchanges would occur in a multifactor cross because there are more markers in a multifactor cross.

The third possibility we will call The Hopeful Hypothesis. Such a hypothesis assumes that HNI is a more or less direct reflection of the mechanism of recombination among closely linked markers, and that, furthermore, this mechanism is intimately related to the mechanism of recombination among loosely linked markers. We can imagine that what is counted as a "single exchange" in a cross involving only distant markers is in reality a complex event which is distributed over a finite region of the genome. We would not expect the results of a cross to be strongly dependent on the detailed structure of the exchange process unless the distances between the markers were of the same order as or smaller than the distance over which the exchange process is distributed. While the hypotheses previously considered cannot be rigorously excluded, the present hypothesis is, in fact, the only one for which there is reasonably cogent experiment support. It includes, but is somewhat broader than, the hypothesis of the "switch region" as conceived by several authors (1,5,6,7). Indeed, we will presently arrive at a prediction which is diametrically opposed to that of the "switch theory" of Freese (7).

3. Heterozygosis.

Further specification of The Hopeful Hypothesis requires some knowledge of the process of recombination among loosely linked markers. In a formal sense, our information about this process comes from the study of phage heterozygotes. If bacteria are mixedly infected with mutant and wild-type phages, most of the progeny phage are homozygous in the sense that upon further growth, each gives rise to a single type--either mutant or wild-type, but not both. A small fraction, however, gives rise to both mutant and wild-type particles and are therefore called heterozygotes. For phage T2 the heterozygous fraction is about 2%; for the closely related phage T4, this fraction is somewhat smaller (1.4%), a result which is probably due to technical difficulties of scoring heterozygotes with this phage. With r-type plaque morphology markers, heterozygous particles yield "mottled" plaques which are distinguishable from both r and r^+ plaques. It was the observation of such mottled plaques that led to the discovery of the phenomenon of heterozygosis in phage T2 by Hershey and Chase (8). These authors described several fundamental features of heterozygosis:

(a) Heterozygosis occurs for all markers with about the same frequency.

(b) Heterozygotes are unstable. A mottled plaque does not contain more than 2% heterozygotes.

(c) When two unlinked markers are used, phage which are heterozygous for one marker are rarely (6%) heterozygous for the other.

(d) When two closely linked markers are used, phage heterozygous for one are often heterozygous for the other. The actual frequency of double heterozygosis varies inversely with the distance between the markers and presumably approaches 100% for exceedingly closely linked markers.

(e) When linked markers are involved, there are always two predominant genotypes in a mottled plaque. In plaques originating from double heterozygotes, the two minority genotypes are presumed to arise secondarily from recombination during growth in the plaque. The structure of the original heterozygote is thus inferred from the predominant genotypes in the plaque.

(f) One of the predominant genotypes in a mottled plaque is always parental; the other may or may not be recombinant.

These results are easily understood by assuming that a phage heterozygote is heterozygous for a short segment of the genome.

The relationship between heterozygosis and recombination was demonstrated by Levinthal (9). Using phage T2 genetically marked with three linked factors, Levinthal showed that phage heterozygous for the middle marker were predominantly recombinant in genotype for the two outside

markers. That they are not all recombinant is easily accounted for on the basis of multiple matings. This finding resulted in the overlap model for the phage heterozygote shown in Figure 1. It is reasonable, of course, to assume that overlaps occur whether or not appropriate genetic markers are present to reveal them as heterozygotes. Overlaps are recombinant for outside markers and, ipso facto, form one mechanism of recombination. Are there other mechanisms? Levinthal calculated that the frequency of overlaps should be sufficient to account for all observed recombination among loosely linked markers. Hence, it can be assumed that recombination among loosely linked markers occurs by means of the overlap mechanism. Similar conclusions were reached by Trautner (10) on the basis of experiments with phage T1.

It should be emphasized that the overlap is a purely formal conception. Of course, a heterozygous region must be diploid in some sense, but it is entirely possible that the entire genome is diploid in the same sense. According to this view, an overlap represents an overlap of "information sources" and would not be physically distinguishable unless it were heterozygous. The most popular physical model for the phage heterozygote has been of this type. Following Levinthal (9), it has usually been assumed that the diploidy of the phage genome reflects the duplex nature of the DNA molecule, and that the heterozygote is struc-

turally a "heteroduplex." It is equally possible, however, that an overlap represents a structural singularity. Doermann and Boehner (11), on the basis of recent experiments, assert that this latter view is correct and that the overlap represents an overlap of material structures. We will not pursue further this question.

4. Relationship of Heterozygosis and Recombination.

To return now to The Hopeful Hypothesis, it is obvious that an overlap represents an exchange which is distributed over a finite region of the genome. Since the distance between what we have termed very closely linked markers is considerably less than the overlap length, about 2% of the particles in the vegetative pool will be heterozygotes of the type which segregate principally the two parental genotypes; i.e., non-recombinant heterozygotes (NRH's). It should be apparent that in order to have a sufficient explanation for HNI, it is only necessary to postulate that an NRH segregates recombinants with an appreciable frequency. Edgar (12) has provided cogent evidence that this is indeed the case. Edgar begins with a lysate of phage T4 containing the progeny from a biparental cross involving two exceedingly closely linked markers. He demonstrates that this lysate contains a small fraction of particles which are not recombinants but which give rise to recombinants upon singly infecting a sensitive bacterial host. The

only type of particle in the lysate which could have this property is the NRH. An NRH would be expected to segregate the two parental types which, in turn, would be expected to produce recombinants just as in a mixedly infected cell. The crucial observation is a quantitative one. The probability that an NRH will give rise to a recombinant is considerably greater than the corresponding probability for a mixedly infected cell. Thus it is concluded that while the principal segregants of an NRH are parental in genotype, there is an appreciable probability that an NRH will segregate recombinants. As Edgar points out, however, mixedly infecting a cell is not quite comparable to infecting it with an NRH. The question of topography again arises (see Part One, Chapter II). The genomes of the phages infecting a mixedly infected cell are far apart compared to the strands of the heterozygote. One can imagine that the heterozygote segregates into the two parental types which are also physically adjacent. One can further imagine that this physical proximity will lead to "incestuous matings" among the segregants, resulting in an excess of recombinants being produced. If our objective is to account for HNI, this question is not necessarily relevant. The NRH's which are formed during the course of a mixed infection would also be expected to segregate parental types which would be physically adjacent and which would undergo incestuous matings. The only precaution

is that the phrase, "segregate recombinants," must be interpreted in a sense sufficiently broad to include the result of incestuous matings.

On the basis of the observations reviewed above, Edgar (12) proposes the following unified mechanism of recombination for all distances. At all distances, a heterozygote (or overlap) is assumed to be an intermediate in the process of recombinant formation. Depending upon the relative position of the overlap region and the markers, there are three distinct ways in which recombinants can be formed in a two-factor cross; these ways are diagrammed in Figure 2. In the first or "direct" route, the overlap lies completely in the region between the markers and no heterozygous intermediate is detectable. In the second or "recombinant heterozygote" (RH) route, one end of the overlap lies between the markers, forming an RH; the RH has a recombinant type as one of its principal segregants. In the third or "NRH" route, both markers lie in the overlap region; the NRH, with some frequency, is able to segregate recombinants--perhaps directly, more likely through an RH intermediate. The relative contribution of each of these routes to the production of recombinants in any given cross will depend upon the distance between the markers. A complete enumeration of the possible routes for the formation of double recombinants in three-factor crosses would be too lengthy to present here; these routes, moreover, would consist of obvious combinations of the three

basic routes.

5. A Critical Test.

The Edgar scheme leads to an unusual prediction. Consider a biparental cross involving a group of very closely linked markers and, in addition, outside markers loosely linked on either side. We select a class of progeny from the cross which are recombinant for the closely linked markers and ask: what is the genotypic composition of this selected class with respect to the outside markers? We reason as follows. The selected recombinants arise as segregation products of heterozygotes. Furthermore, heterozygotes are predominantly recombinant for outside markers. Therefore, we would predict that the selected recombinants will be predominantly recombinant for the outside markers, irrespective of how many exchanges are required to produce the selected recombinant class. In the cross schematically shown in Figure 3, three closely linked rII markers (13) are flanked by the markers A and B in coupling. According to the Edgar scheme, r^+ recombinants are formed from an intermediate NRH which is either A+ or +B in genotype. Secondary recombinational events would produce AB and ++ as minority genotypes. Translated into the language of classical genetics, this means that triple exchanges would be more frequent than double (or quadruple) exchanges--a most unusual prediction

indeed. The classical rules lead, of course, to the prediction that AB (double exchange) will be more frequent than A+ or +B (triple exchange) which in turn will be more frequent than ++ (quadruple exchange). Thus the prediction of the Edgar scheme is far from trivial. It differs so strongly from the classical picture that we thought that crosses of the type shown in Figure 3 would provide a critical test of our current ideas about the relationship of heterozygosis and recombination. Experiments along these lines will be presented in the sequel. Before proceeding further, however, let us state the result. The principal result of Part Two of this Thesis will be that the segregation pattern of outside markers is always in accordance with the simple rules of classical genetics. Hence we will conclude that something is wrong with our current ideas.

II. MATERIALS AND METHODS

1. Bacterial Strains.

Several strains of Escherichia coli were used. Strain B was used as a host for crosses. Strain S/6 (1^4) was used as the standard non-selective indicator strain. Strain K is the K-12 derivative 112-12 (15) lysogenized with phage lambda by J. J. Weigle. Strain F, a hybrid between K-12 and B, was given to us by P. Fredericq (16) and lysogenized with lambda. Both F and K support the growth of wild-type T^4 but do not support the growth of rII mutants (13) and were used as selective indicator strains.

2. Phage Strains.

The wild-type strain T^4D and various mutant derivatives were employed.

T^4D rII mutants were isolated and mapped in our laboratory (3): rEDa^{4}1, rEDb^{4}5, rEDb^{4}6, rEDb^{4}8. (In the sequel, only the numerical part of the designation will be used.)

T^4D ac^{4}1 was isolated by plating a high titer stock of T^4D on acriflavine plates. (See below.)

T^4D tu^{4}1 is described by Doermann and Hill (13).

T^4D ti originated in the following manner. A "partial revertant" of r^{4}1 was selected by plating a high

titer stock on strain F. This revertant, designated r41su4, gives r-type plaques on S/6, small plaques with sharp margins on F (with good efficiency), and "pinpoint" plaques on K. Similar partial revertants had been shown by Feynman (17) to be due to closely linked suppressor mutations; he further showed that the suppressor was itself a standard rII mutant. Our isolate, r41su4, was backcrossed to wild-type T4D; r plaques were picked and tested for ability to grow on K. Any recombinants would not be expected to grow on K. Out of 800 plaques tested, all grew on K. Thus, at a confidence level of 99%, the suppressor must be within 0.4 map units of r41. For our purpose this justifies treating the suppressed r as if it were a single mutant, designated "tiny" (ti) since it gives tiny plaques on strain F.

3. Media.

H-Broth: Bacto Nutrient Broth, 8 gm.; Bacto Peptone, 5 gm.; sodium chloride, 5 gm.; glucose, 1 gm.; tap distilled water, 1ℓ. H-Broth was used as a liquid nutrient medium in all experiments.

EHA-bottom layer: Bacto agar, 10 gm.; Bacto Tryptone, 13 gm.; sodium chloride, 8 gm.; sodium citrate (dihydrate), 2 gm.; glucose, 1.3 gm.; tap distilled water, 1ℓ. Approximately 40 ml. per plate of EHA-bottom layer is used.

EHA-top layer: Bacto agar, 6.5 gm.; Bacto Tryptone, 13 gm.; sodium chloride, 8 gm.; sodium citrate (dihydrate), 2 gm.; glucose, 3 gm.; tap distilled water, 1 ℓ. Approximately 2 ml. of EHA-top layer is used in the agar overlay technique (18) for plating.

Acriflavine supplemented plates contained 0.25 mg./ℓ. "acriflavine neutral, NF IX" (Nutritional Biochemicals Corporation) in both bottom and top layers.

4. Preparation of Phage Stocks.

All phage stocks were clonally derived from single plaques. Phage of the desired genotype are plated on S/6, and the plates are incubated at 30° C. for four hours. A plug of agar containing a single well isolated plaque is removed with a capillary tube and transferred to about 25 ml. of a growing culture of bacteria in H-Broth. The growing culture is obtained by diluting a saturated culture of S/6 1:500 or 1:1000 into H-Broth and incubating for 2 1/2 hours at 30° C. with aeration. After addition of phage, incubation is continued until lysis. The lysate is shaken with chloroform, centrifuged in the cold (ca. 4,000 x g for 20 min.), and then filtered through a Mandler candle. Stocks prepared in this way are usually stable indefinitely in the refrigerator. They can be used for high multiplicities of infection (usually up to 30 phage/cell) without lysis from without (19). They are

generally of sufficient genetic homogeneity for use in crosses.

5. Plating Conditions.

Cultures of the indicator strains in the late exponential phase of growth were used for plating. A saturated culture is diluted 1:100 into H-Broth, incubated for 2 1/2 hours at 30° C. with aeration, centrifuged in the cold, and resuspended in 1/10 volume of fresh broth. Such cultures may be used for several days if kept in the cold. Two or three drops were used for plating.

S/6 plates were incubated at 30° C. and K plates at 37° C. for 18 to 24 hours. For routine assays, F plates were incubated at 37° C. for 18-24 hours. When cross progeny were plated for scoring of genotypes, another regime was used in addition: incubation at 37° C. for one hour followed by incubation at 25° C. for two days. At the lower temperature, the ti plaques are much smaller and more easily distinguishable; the preliminary incubation is necessary, however, to improve the efficiency of plating and uniformity of size. The distinguishability of the tu character is also better at the lower temperature, at least in a ti^+ background. Unfortunately, plaque size is not reproducible from day to day, and ti plaques were often too small to permit scoring of the tu character as well, particularly on acriflavine plates. Accordingly, both

incubation regimes were usually used. Appropriate control mixtures were used to test the reliability of scoring each day. The efficiency of plating of each genotype was also determined each day by plating appropriate mixtures. When the frequency of $\underline{r}^+$ in the progeny of a cross is low, the possibility exists that markers from parental $\underline{r}$ phage are "rescued" into $\underline{r}^+$ phage from accidental mixed infections occurring on the plate. Hence, control platings in which the parental mixture was also added were performed each day. It was always possible to plate a sufficiently small amount of the cross progeny for this sort of contamination to be indetectable.

6. Standard Cross Procedure.

Our procedure follows, in general, that of Chase and Doermann (1). A saturated culture of B bacteria is diluted 1:1000 into H-Broth, incubated for 2 1/2 hours at 30° C. with aeration, centrifuged in the cold, and resuspended in a small volume of fresh broth. The cells are counted in a Petroff-Hausser counting chamber under the phase contrast microscope and diluted to a concentration of 4×10^8 bacteria/ml. A few minutes before the addition of phage, sufficient potassium cyanide is added to make the bacterial suspension 0.004 M. Equal volumes of the parental phage mixture (usually containing about 6×10^9 total phage/ml.) and bacterial suspension are

mixed. Ten minutes is allowed for adsorption to take place, and during this interval the mixture is agitated by gentle aeration at 30° C. After the adsorption period, the infected culture is diluted at least 10^4-fold into a growth tube (GT) containing H-Broth prewarmed to 30° C. An aliquot of the GT is plated immediately for total infective centers. To another aliquot, chloroform is added to kill infected cells. Unadsorbed phage are enumerated by plating the chloroform treated aliquot, and the number of infected cells is obtained by difference. Usually more than 99% of the parental phage are adsorbed. The GT is aerated at 30° C. for 90 minutes, at which time chloroform is added to sterilize the tube. Progeny phage are counted and scored for genotype by diluting and plating under appropriate conditions.

7. Premature Lysis Procedure.

When it is desired to examine the progeny produced before the normal lysis time, the standard cross procedure cannot always be used since insufficient numbers of progeny phage may be present to yield a statistically reliable sample of recombinants. The simple expedient of diluting the infected cells less after adsorption is not satisfactory due to residual cyanide. When the adsorption mixture is diluted only 10^3-fold, there is an appreciable delay in the appearance of the first intracellular phage. This is

somewhat puzzling in view of the finding of Benzer and Jacob (20) that five-fold higher concentrations of cyanide do not appreciably delay the earliest steps of intracellular development as measured by sensitivity to ultraviolet light. One could speculate that maturation is a particularly cyanide-sensitive process, but we have not pursued the matter further.

To avoid the residual cyanide effect, the adsorption mixture is diluted 40-fold into ice-cold H-Broth and centrifuged at 0-4° C. The supernatant is decanted, and the upper part of the tube is wiped with sterile cotton swabs. During this procedure, particular care is taken to avoid warming the infected cells in the pellet. The pellet is resuspended in about 2 ml. of cold H-Broth. Development begins when the resuspended cells are transferred to a growth tube (GT) containing at least 100 ml. of H-Broth prewarmed to 30° C. This concentrated GT is continuously aerated, and infected cells and unadsorbed phage are measured in the usual way. It should be noted that only about 1/3 of the infected cells are recovered. Most of the loss presumably arises from centrifuging a dilute suspension in a fairly large tube. It should also be noted that despite the fact that no more than 1% of the original adsorption medium could be carried over into the concentrated GT, the ratio of apparently unadsorbed phage to infected cells changes very little. This

presumably means that desorbed phage make an appreciable contribution to the unadsorbed phage. Hence, correction of the progeny yields for unadsorbed phage is likely to be rather inaccurate, and we have disregarded all samples for which this correction is more than 10%.

Premature lysates are obtained by adding chloroform to aliquots of the concentrated growth tube at various times. When it is desired to store undiluted lysates, they are filtered.

III. EXPERIMENTAL RESULTS

1. Introduction.

At the beginning of this study, we were faced with two problems. First, no suitable set of markers was available. Furthermore, plaque morphology on the bacterial strains usually used for the selection of r^+ recombinants was not sufficiently uniform for our purposes. The problem of a suitable selective indicator strain was solved by the use of strain F. Wild-type T4 phage produce reasonably uniform plaques on this host with an efficiency approaching unity; rII mutants produce no plaques. Considerable effort was expended in isolating new markers. The linkage relations of the markers finally chosen for use in the present experiments are shown in Figure 4. With the exception of tu41, these markers were all isolated by us; a detailed description was given in Chapter II. The map distances given in Figure 4 are values obtained after spontaneous lysis.

In order to minimize the effects of multiple matings, all of the crosses to be reported below were executed according to the premature lysis procedure described in Chapter II. Genotypic frequencies were determined in lysates in which the average phage yield was 7.5 to 15 phage per cell. The lysate was plated on strain F. One set of plates was supplemented with acriflavine, and

another set of plates was not. On both sets of plates, plaques were scored for the ti and tu characters by inspection of plaque morphology. The frequencies of genotypes bearing the ac allele are thus determined directly, while the frequency of genotypes bearing the ac^+ allele are determined by difference. As explained in Chapter II, small efficiency of plating (e.o.p.) corrections were usually necessary. When the selected r^+ recombinants are predominantly ac, small volumetric errors and errors in the determination of e.o.p. are amplified by the subtraction of one large number from another, and a high order of accuracy is not to be expected. Little difficulty was experienced in scoring plaque types on standard plates. On acriflavine plates, the scoring is somewhat difficult, particularly when both the ti and the tu characters must be scored simultaneously. Accordingly, the tu marker was not included in all crosses. The total number of r^+ recombinant plaques scored ranged from 4,000 to 10,000 in various crosses (except for cross #42 in which about 2,700 plaques were scored). Hence, statistical sampling is not the limiting source of error in any frequency exceeding a few per cent.

Lysates were also plated on the non-selective indicator strain S/6 on standard plates and on acriflavine plates. The ratio of plaques on F to plaques on S/6 gives the fraction of r^+ recombinants. The ti marker cannot be

scored in a r background on S/6. Hence the only additional information obtained in the crosses involving closely linked rII markers is the frequency of exchanges in the ac-tu interval for crosses in which the tu marker is present. One cross involved only a single rII marker; in this case more information was obtained from the S/6 plating.

2. Crosses Involving Three Closely Linked rII Markers.

We have performed four crosses, each involving three closely linked rII markers arranged in such a fashion that two exchanges are required to produce the selected r^+ recombinant type. Each cross involves a different one of the four possible arrangements of mutant and wild-type alleles at the ti and ac loci. The tu marker is included in only two of the crosses. Genotypic frequencies in the progeny from each cross are given in Table 1. Inspection of Table 1 reveals that qualitatively normal linkage relations are maintained in all crosses. It will be noticed that there is considerable variation from cross to cross and that this variation is not random. The variability and other quantitative considerations will be discussed in later sections.

3. Crosses Involving Two Closely Linked rII Markers.

We have performed two crosses, each involving two closely linked rII markers (r^{46} and r^{48}). In these

crosses, a single exchange is sufficient to produce the selected r^+ recombinant type. The ti and ac markers are included in both crosses; the tu marker is included in only one. Genotypic frequencies in the progeny from each cross are given in Table 2. Inspection of Table 2 reveals that qualitatively normal linkage relations are again maintained. Quantitative considerations are deferred until later sections.

4. A Cross Involving Only One rII Marker.

We have performed one cross involving only one rII marker (r^{46}). This cross differs from the crosses previously considered in that by plating on strain F, we select the parental r^+ allele, rather than an r^+ recombinant type. The results are given in Table 3. It is apparent that a few of the genotypes can be determined independently by plating on strain S/6, and we shall have occasion to refer to such results later. The frequencies given in Table 3, however, are based solely upon the platings on F; thus, the data of Table 3 are directly comparable to those obtained from the selective crosses and given in Tables 1 and 2.

5. Heterozygosis.

Let us return to the crosses involving three closely linked rII markers. In this section, we will ignore the

tu marker and concern ourselves solely with the markers adjacent to the r region. For each of these crosses, we calculate from the data in Table 1 the frequency of exchanges in the ti-r interval (interval I), the r-ac interval (interval II), and the frequency of simultaneous exchanges in I and II. These exchange frequencies are given in Table 4. The average frequencies of exchange are quite similar for the two intervals. As noted previously, however, there is considerable variation from cross to cross. When the data from Table 4 are rearranged as in Table 5, it is seen that a considerable part of the variation is not random. When an exchange in interval I (or II) results in a recombinant type which bears the mutant allele at the ti (or ac) locus, the apparent frequency of exchange is considerably smaller than when the exchange results in the recombinant being wild-type. The effect is still more drastic when double exchanges are considered. The double recombinant may bear 0, 1, or 2 mutant alleles, and the apparent exchange frequency decreases in the same order.

An asymmetry between mutant and wild-type alleles is shown by one of the crosses (cross #47) involving two closely linked rII markers (Table 2), but not by the other cross. That is, one of these crosses is not only inconsistent with the observations made in the preceding paragraph but also with the other cross of the same type.

In view of this, we will not further consider the crosses shown in Table 2.

The asymmetry between mutant and wild-type alleles is not due to selection or to different efficiencies of plating (e.o.p.'s), at least not in the ordinary sense. When *ac* and *ti* are backcrossed to wild-type, no selective effects are demonstrable. Furthermore, the data have already been corrected for e.o.p. differences as determined by plating control mixtures of different genotypes. Special selective or e.o.p. effects which are only demonstrable when recombinants for closely linked markers are selected cannot be excluded.

A more reasonable explanation in terms of known phenomena could be given, however, if a significant fraction of the selected r^+ recombinants were heterozygous for *ti* or *ac* (or both). It is apparent that an adequate explanation of the data in terms of heterozygosis requires that particles *heterozygous* for *ti* or *ac* be phenotypically wild-type. A direct determination of the phenotype of *ac*- or *ti*- heterozygotes is not feasible. However, if cells are mixedly infected with high equal multiplicities of *ti* and ti^+ phage and plated on strain F, more than 99% of the plaques are ti^+ in morphology. Similarly, of cells mixedly infected with high equal multiplicities of *ac* and ac^+, less than 1% give rise to plaques on acriflavine plates. Hence, *ti* and *ac* are both recessive in the sense

defined above, and it is reasonable to assume that heterozygotes for either marker will be phenotypically wild-type. The recessiveness of ti is readily explicable in terms of selective forces which are operative during the course of plaque formation on strain F. It is, in fact, such selection that makes the direct determination of the phenotype of a ti-heterozygote impossible. It is more difficult to understand why ac, an allele for resistance, should be recessive to ac^+, an allele for sensitivity, but the effect is quite real and reproducible (21).

In order to test the "heterozygote hypothesis," as well as for its intrinsic interest, let us attempt to calculate the frequencies of the various classes of heterozygotes which must be postulated in order to account for our results. At either locus (ac or ti), a given particle may be homozygous for allele 1 (see Table 1), homozygous for allele 2, or heterozygous. Thus, there are three states at each locus or nine classes in all. We shall assume strict symmetry, which reduces the number of different frequencies to be found to six. Since the frequencies of all classes must sum to unity, only five of the frequencies are independent. The five quantities, A, B, C, D, and E, in Table 5 are thus sufficient to calculate the frequencies of all nine classes. For example, C, the frequency of the double recombinant when that type is a double mutant (ti ac), must be the frequency of the "pure," non-hetero-

zygous double recombinant. We shall not go through the algebra involved in calculating the other quantities here. The results and numerical calculations are given in Table 6.

Turning now to Table 6, we note that the "heterozygote hypothesis" is reasonable in the sense that the calculated frequencies of all classes of particles are positive. The total frequency of heterozygotes is calculated to be 20.4%, a quite significant fraction. Aside from this, no great reliance should be placed on the numerical values. With this warning, however, we proceed to note a few points. Most of the heterozygotes are heterozygous for only one marker. Indeed, the frequency of double heterozygosis (1.0%) is close to the square of the frequency of heterozygosis for an individual marker (10.7%), and in this sense, heterozygosis occurs independently for each locus. The frequency of heterozygosis among our selected recombinants is considerably greater than the frequency (2%) observed in the whole population. In view of the observations of Edgar (22) that a considerable fraction of the phenotypically $\underline{r}^+$ are actually heterozygotes, it seems reasonable to assume that the particles heterozygous for an outside marker are also heterozygous for at least one of the rII markers. That is, the heterozygotes observed here are actually quite long. The only relevant data on the distribution of lengths of overlap regions are those of

Doermann (11). From Doermann's data, we estimate that, of particles heterozygous for one of the rII markers, 15-20% would also be heterozygous for one of the outside markers. If we assume that the distribution of overlap regions in our case is similar, then at least 50% of the selected r^+ recombinants are actually heterozygotes. Alternatively, we could assume that less than 50% of the selected r^+ recombinants are heterozygotes, but these heterozygotes would then be longer, on the average, than in the entire population.

It is also possible to calculate genotypic frequencies for the separate "strands" of the heterozygotes. Such calculations show that qualitatively normal linkage relations are maintained for the heterozygous classes as well as for the pure types. The exchange frequencies calculated for the heterozygous strands are somewhat greater than for the non-heterozygous classes. In Chapter IV, we shall seriously consider a hypothesis which in the present context amounts to saying that the pure types arise from heterozygotes. The numerical calculations, for what they are worth, do not contradict this hypothesis.

An effect of heterozygosis is also evident in the cross in which the selected r^+ class is not necessarily recombinant (cross #50, Table 3). The apparent frequencies of exchange are given in Table 7. The exchange frequencies in this cross are very much lower than in the crosses pre-

viously considered. Hence, we would expect heterozygotes occurring with the normal frequency to influence the results. For this cross, it is also possible to determine the frequency of exchange in interval I by plating on indicator strain S/6. On strain S/6, selection against the ti allele is not strong, and ti-heterozygotes (usually) give rise to mottled plaques. The frequency of exchange in interval I determined by plating on S/6 is 0.0274; this value is considerably smaller than the value of 0.0348 obtained by plating on F and is quite close to the frequency of exchange in interval II determined by plating on F, 0.0258.

6. Interference.

In this section, we will consider several types of interference relationships. Before proceeding further, however, a few cautionary notes should be introduced. In view of the conclusion that a significant fraction of the r^+ recombinant progeny are heterozygous for the outside markers, the meaningfulness of any interference calculation is open to question. We shall, wherever possible, use average values to eliminate some of the biases due to heterozygotes. All of our data concern cross progeny obtained by premature lysis of the infected cells. With our procedure, it is impossible each day to lyse the cells at exactly the same point in the phage life cycle. Thus the data are non-homogeneous to this extent. Furthermore, as noted previously, some of the procedures used in order

to obtain frequency data are not intrinsically very accurate. Therefore, we should not expect to be able to make fine comparisons and should restrict ourselves to rather gross effects.

First, let us consider interference in the production of $\underline{r}^+$ recombinants, without reference to the outside markers. In Table 8, we give the frequency of $\underline{r}^+$ recombinants in each of our crosses. The data in Table 8 allow us to estimate the frequency of double exchanges and the frequency of exchanges in one of the two intervals. No comparable data are available for the other interval. Since the frequencies of recombination are essentially the same in the two intervals at the time of normal lysis (see Figure 4), it is reasonable to assume that they are also the same at the lysis times used here. Thus, we calculate in Table 8 that the frequency of double exchanges exceeds random expectation by a factor of slightly over 50. This factor is about 3 to 4 times the interference index obtained by Chase and Doermann (1) for markers at comparable distances. The discrepancy is, of course, explicable in terms of different lysis times in the two cases. It is, in fact, generally true that the interference index decreases with increasing time of lysis (3). It is perhaps more instructive to look at the absolute values rather than the interference index. Given an exchange in one interval, there is a probability of about

10% that there will be an exchange in the other interval.

Let us now consider interference in the intervals adjacent to the rII markers. This can be looked at in two ways. First, we can compare the frequency of exchange in interval I (or II) among the selected r^+ recombinant class with the corresponding frequency of exchange among all particles. The latter must be determined in a separate cross (cross #50). Furthermore, we can compare the frequency of double exchanges in interval I and interval II with a random expectation based upon the frequencies of exchanges observed in the separate intervals among the r^+ recombinants. The first comparison gives us a measure of interference due to selection of r^+ recombinants, while the second gives a measure of additional interference within this class. The relevant data and calculations are shown in Table 9. The data in Table 9 are based upon the crosses involving three closely linked rII markers together with data from the cross with selection for an r^+ parental allele for comparison. We have assumed that the markers are symmetrically placed and have averaged all individual frequencies. An examination of Table 9 reveals that the frequency of exchange for a single interval (I or II) is over 8-fold greater among the r^+ recombinant class than among the entire progeny. On the other hand, there is essentially no additional interference for double exchanges within the r^+ recombinant class.

Thus far we have neglected the tu marker. We will refer to the ac-tu interval as interval III. The frequency of exchanges in interval III (unlike in intervals I and II) may be determined over the entire progeny by plating on S/6 on acriflavine plates and scoring the tu character. In Table 10, we present the frequency of exchanges in interval III in both the selected r^+ class and in the entire population; all crosses involving the tu marker are included in this table. We first note that for the cross in which the selected r^+ class is not necessarily recombinant (cross #50), the two methods of determination agree quite well. If we consider the average of the other crosses (in which the selected r^+ progeny are recombinants), the two determinations are not only similar to each other, but to the former determinations as well.* That is, there is no interference. This is a rather surprising and intriguing result. As discussed in Part One of this Thesis, it has invariably been true that when any recombinant class is selected, the frequency of any other recombinant class is greater than for the population as a whole. This was explained in terms of heterogeneity in "mating experience." In view of this, our finding that there is no interference, if true, would really mean

---

*Taken at face value, there would appear to be a difference between the results when one exchange and when two exchanges are selected. Whether this is significant or not cannot be said.

that there is positive interference in a single "mating." We will not present further interference calculations in detail. In the cross in which the selected r^+ type is not necessarily recombinant, the expected negative interference for exchanges in interval III is demonstrable when considering any recombinant class. In the selective crosses, there is a comparable degree of negative interference for exchanges in interval III among particles recombinant in the adjacent interval II. Among particles recombinant in interval I, there is, if anything, positive interference. The significance of these findings is by no means clear.

IV. DISCUSSION

The new experimental results reported here concern the segregation pattern of loosely linked outside markers within a selected class of particles which are recombinant for closely linked internal markers. We can summarize our observations as follows:

(a) The outside markers maintain qualitatively normal linkage relations with the internal markers and with each other. This is true whether one or two exchanges are required to produce the selected recombinant class.

(b) The region of HNI extends into the intervals adjacent to those in which the exchanges are selected but does not extend into a non-adjacent interval.

(c) Multiple exchanges within the selected class occur with little, if any, additional interference. That is, given the frequencies of exchanges for the separate intervals within the selected class, the frequency of multiple exchanges is essentially equal to the random expectation.

(d) A significant fraction of the selected recombinant class is heterozygous for at least the adjacent outside markers. The evidence upon which this conclusion is based is indirect and is, to this extent, weaker than the preceding observations.

The fact that, aside from the complications of negative interference and heterozygosis, the segregation pattern of outside markers is normal would, of itself, require

little comment. In view of the previously discussed (Chapter I) current notions concerning the mechanism of recombination in phage, the "complications" are to be expected, while the "normal" result leads us to a paradox. We are now faced with three apparently well founded assertions:

(a) According to Edgar (12), recombinants for very closely linked markers arise as segregation products of heterozygotes--either directly or as a result of incestuous matings among the primary segregation products of heterozygotes.

(b) According to Levinthal (9) and to Trautner (10), heterozygotes are recombinant for outside markers.

(c) According to our observations, recombinants for closely linked markers are not necessarily recombinant for outside markers.

We do not think that all of these assertions can be strictly correct.

We do not pretend to be able to resolve the paradox at the present time. We can only examine it a little more closely and speculate about the direction from which an eventual resolution might come. It is possible, of course, that the paradox is the result of an error of fact, but we consider this to be unlikely. If there is no error of fact, there must be an error of interpretation.

First of all, it should be noted that we have tacitly assumed that there is only one type of heterozygote. Next, we consider the assertions individually. In reality, the three paradoxical assertions are not highly derivative. They follow very closely from experimental observations, and were we not faced with an apparent paradox, we would be tempted to refer to them as "facts" rather than as "assertions." The factual basis for Edgar's assertion is that an NRH, upon singly infecting a host cell, segregates recombinants with a frequency great enough to account for the production of recombinants for closely linked markers in a mixedly infected cell. It is assumed that an NRH which arises in a mixedly infected cell behaves in an identical manner. At first sight, the other two assertions might appear to be direct statements of the experimental observations, but in fact they involve a similar type of assumption. The factual basis for the Levinthal-Trautner assertion is that the progeny of heterozygotes (in mottled plaques) are predominantly recombinant for outside markers. It is assumed that the heterozygous particles (which give rise to mottled plaques) are themselves recombinant. A similar caveat is applicable to our own assertion. We do not directly examine recombinants for closely linked markers; we examine the progeny of such recombinants at some remove. These would seem to be the important assumptions. There may be others.

If any of the above mentioned assumptions were not strictly correct, it would be easy to resolve the paradox. If there were two types of heterozygotes, for instance, then Levinthal and Trautner could be studying one type and Edgar the other type, and there is no paradox. If an NRH which arises during the course of a mixed infection behaves somewhat differently from a singly infecting NRH, there are several possibilities. One of the more likely ones involves the previously discussed (Chapter I) notions of topography and incestuous matings. We might imagine that an NRH which arose during the course of a mixed infection produces recombinants as a result of an incestuous mating with one (or both) of the parental types which gave rise to the NRH. This could not happen, of course, with a singly infecting NRH.

We choose to speculate, however, that the fallacy lies in reasoning that a heterozygote is recombinant for outside markers because its offspring are so recombinant. If the primary heterozygote is not recombinant for the outside markers, what is its genotype? We would not wish to argue that it is parental in genotype; hence, it must be heterozygous for one or both of the outside markers as well. We have already pointed out that it is possible that the entire phage genome is diploid in the same sense that a heterozygous region must be diploid. We now note

that it is not even necessary to assume that the major portion of the genome is homozygous. Let us consider a rather farfetched analogy. The ciliate, Paramecium aurelia, is diploid. A regular feature of the life cycle of this organism is a type of internal mating known as autogamy (23). The first step in autogamy is meiosis. There are two nuclei which give rise to eight meiotic products; seven of these degenerate. The surviving meiotic product undergoes a mitotic division, and this is followed by a fusion of the two daughter nuclei. The net result is the production of a completely homozygous diploid individual. Now if, for some reason or other, it were impossible to observe the immediate exconjugants from a mating, and genetic studies had to be performed by studying only large clones arising from individual exconjugants, it would be concluded that P. aurelia was a haploid organism! The moral of this story should be obvious. We might postulate that phage T4 is a diploid organism but that it replicates vegetatively in such a fashion that all of its progeny usually bear the same allele at any given genetic site. Over a short region of the genome, both alleles may appear in the progeny, thus enabling us to recognize that the progenitor was heterozygous. In this light, our finding that a significant fraction of the selected recombinants are heterozygous

for the outside markers assumes a new importance. Compatible with the diploid postulate is a recent result of Doermann and Boehner (11). They have analysed the phage present in mottled plaques arising from particles heterozygous for several fairly closely linked markers. They find that, in a given plaque, the frequency of the mutant allele in the plaque is different for different markers. In particular, a given allele may be present in very low frequency. This suggests that the original particle was also heterozygous for other markers which, however, did not appear at all in the progeny. We should note that Doermann and Boehner interpret their results in quite a different fashion (see Chapter I). In any event, it is likely that a resolution of our paradox will emerge from further studies on heterozygotes.

TABLE 1

Crosses Involving Three Closely Linked rII Markers

$$\frac{1 \qquad + \; r \; + \qquad 1 \qquad 1}{\text{X}}$$

$$\frac{2 \qquad r \; + \; r \qquad 2 \qquad 2}{}$$

Cross no.	Genotype of +r+ parent (111)	Genotype of r+r parent (222)	Frequency, among selected r^+ recombinant class, of genotype: (111) (112)	(211) (212)	(122) (121)	(222) (221)
32	ti + +	+ ac tu	0.477 0.020 → 0.497	0.266 0.013 → 0.279	0.134 0.031 → 0.165	0.052 0.0072* → 0.059
33	+ ac +	ti + +	0.528	0.129	0.276	0.068
40	+ + tu	ti ac +	0.550 0.074 → 0.624	0.137 0.021 → 0.158	0.167 0.013* → 0.180	0.034 0.0040* → 0.038
42	ti ac +	+ + +	0.518	0.189	0.193	0.100

*Frequency based upon a plaque-count of less than 100

TABLE 2

Crosses Involving Two Closely Linked rII Markers

$$\frac{1 \qquad r + \qquad 1 \qquad 1}{\text{X}}$$

$$\frac{2 \qquad + r \qquad 2 \qquad 2}{}$$

Cross no.	Genotype of r+ parent (111)	Genotype of +r parent (222)	Frequency, among selected r^+ recombinant class, of genotype: (211)	(212)	(111)	(112)	(222)	(221)	(122)	(121)
47	+ + tu	ti ac +	0.469	0.043	0.230	0.045	0.140	0.022	0.041	0.0100*
			0.512		0.275		0.162		0.051	
49	ti ac +	+ + +	0.571		0.214		0.191		0.024	

*Frequency based upon a plaque-count of less than 100

TABLE 3

A Cross Involving Only One rII Marker

1	+	1	1

X

2	r	2	2

Cross no.	Genotype of r^+ parent (111)	Genotype of r parent (222)	Frequency (%), in selected class bearing r^+ allele, of genotype:							
			(111)	(112)	(211)	(212)	(122)	(121)	(222)	(221)
50	ti + tu	+ ac +	86.407	7.896	2.659	0.461*	1.800	0.416	0.318	0.054*
			94.30		3.12		2.22		0.36	

*Frequency based upon plaque-count of less than 100

TABLE 4

Exchange Frequencies for Intervals I and II

Cross no.	Frequency of Exchanges in Interval I (ti-r)	Interval II (r-ac)	Intervals I and II
32	0.338 (B)	0.224 (A)	0.059 (D)
33	0.197 (A)	0.344 (B)	0.068 (D)
40	0.196 (A)	0.218 (A)	0.038 (C)
42	0.289 (B)	0.293 (B)	0.100 (E)
Average	0.255	0.270	0.066

NOTES:

(1) Letter in parentheses beside each frequency is used for reference in text and in Table 5.

(2) Data in this table are based upon Table 1. Reading from left to right, the three columns of exchange frequencies are the frequencies of genotypes 21· + 22·, 12· + 22·, and 22·, respectively. The dot (·) in the genotype formulae indicates that the tu allele is disregarded (e.g., 21· = 221 + 222).

TABLE 5

Asymmetry in Exchange Frequencies for Intervals I and II

Apparent Frequency of Exchanges in				
Interval I or Interval II when recombinant type is:		Interval I and Interval II when recombinant type is:		
Mutant (A)	Wild-Type (B)	Double Mutant (C)	Single Mutant (D)	Wild-Type (E)
0.224	0.338	0.038	0.059	0.100
0.197	0.344		0.068	
0.196	0.289			
0.218	0.293			
0.209 (Average)	0.316 (Average)	0.038	0.064	0.100

NOTE:

Data in this table are taken from Table 4. Listed under column A, for instance, are all of the values marked "(A)" in Table 4.

TABLE 6

Calculation of the Frequencies of

Heterozygous and Pure Types

	Allele(s) present at locus		Frequency:	
	ti	ac	Formula	Calculated value
Doubly heterozygous class	1,2	1,2	C-2D +E	0.010
	1,2	1	-A +B +D -E	0.071
Singly heterozygous classes	1	1,2	-A +B +D -E	0.071
	1,2	2	-C +D	0.026
	2	1,2	-C +D	0.026
	1	1	1 -2B +E	0.468
Non-heterozygous classes	1	2	A -D	0.145
	2	1	A -D	0.145
	2	2	C	0.038
			1	1.000

NOTES:

(1) Numerical designations for alleles present at ti and ac loci refer to Table 1.

(2) Letters in formulae for calculation of frequencies refer to Table 5. Average values from Table 5 were used for numerical calculations.

TABLE 7

Exchange Frequencies for Intervals I and II

Exchanges in interval	Frequency	Number of mutant alleles in recombinant type
I	0.0348	0
II	0.0258	1
I and II	0.0036	1

NOTE:

Data in this table are taken from Table 3; they are all based upon platings on indicator strain F.

TABLE 8

Negative Interference in the Production of $\underline{r}^+$ Recombinants

Exchanges selected in interval(s)	Cross no.	Frequency of exchanges	Random expectation	Interference index
	47	1.98×10^{-3}		
$\underline{r}^{46}$-$\underline{r}^{48}$	49	1.84×10^{-3}		
	average	1.91×10^{-3}		
$\underline{r}^{45}$-$\underline{r}^{46}$	Assumed to be equal to interval $\underline{r}^{46}$-$\underline{r}^{48}$ (see text)			
	32	2.15×10^{-4}		
$\underline{r}^{46}$-$\underline{r}^{48}$	33	1.87×10^{-4}		
and	40	1.85×10^{-4}		
$\underline{r}^{45}$-$\underline{r}^{46}$	42	1.93×10^{-4}		
	average	1.95×10^{-4}	3.65×10^{-6}	53.4

NOTES:

(1) In each cross, the frequency of exchanges is twice the observed frequency of $\underline{r}^+$ recombinants in order to take into account the reciprocal recombinant class.

(2) The random expectation for the frequency of double exchanges is the product of the exchange frequencies in the single intervals.

(3) The interference index is the ratio of the observed frequency of double exchanges to the random expectation.

TABLE 9

Negative Interference for Exchanges in Intervals I and II

	Frequency of exchanges when selected r^+ class is		
	Parental allele (a)	Doubly recombinant (b)	Interference index (I_2)
Single interval (I or II)	0.0303	0.262	8.6
Both intervals (I and II)	0.0036	0.066	
Random expectation	0.00092	0.069	
Interference index, I_1	3.9	0.96	

NOTES:

(1) "Random expectation" for each column is the square of the frequency for the corresponding single interval.

(2) Interference index, I_1, is the ratio of the observed frequency of exchanges in both intervals to the corresponding random expectation.

(3) Interference index, I_2, is the ratio of the frequency in column (b) to that in column (a).

(4) Data in column (a) are based on Table 7. Data in column (b) are based on Table 4. Average values are used in both cases.

TABLE 10

Exchange Frequencies for Interval III

Cross no.	No. of selected exchanges	Frequency of Exchanges in Interval III: Selected class	All classes
32	2	0.071	0.097
40	2	0.076	0.080
47	1	0.117	0.081
Average of above		0.088	0.086
50	0	0.087	0.084

NOTES:

(1) The "selected class" data are based on Tables 1, 2, and 3; they are derived from platings on strain F.

(2) The "all classes" data are derived from platings on strain S/6.

REFERENCES

1. Chase, M., and A. H. Doermann, 1958. High negative interference over short segments of the genetic structure of bacteriophage T4. Genetics, 43: 332-353.

2. Edgar, R. S., and C. M. Steinberg, 1958. On the origin of high negative interference over short segments of the genetic structure of bacteriophage T4. Virology, 6: 115-128.

3. Steinberg, C. M., and R. S. Edgar. Unpublished observations.

4. St. Lawrence, P., and D. M. Bonner, 1957. Gene conversion and problems of allelism. The Chemical Basis of Heredity, W. D. McElroy and B. Glass, eds., pp. 114-122. The Johns Hopkins Press, Baltimore.

5. Hershey, A. D., 1958. The production of recombinants in phage crosses. Cold Spring Harbor Symposia on Quantiative Biology, XXIII: 19-46.

6. De Serres, F. J., 1958. Recombination and interference in the ad-3 region of Neurospora crassa. Cold Spring Harvor Symposia on Quantitative Biology, XXIII: 111-118.

7. Freese, E., 1957. The correlation effect for a histidine locus of Neurospora crassa. Genetics, 42: 671-684.

8. Hershey, A. D., and M. Chase, 1951. Genetic recombination and heterozygosis in bacteriophage. Cold Spring Harbor Symposia on Quantitative Biology, XVI: 471-479.

9. Levinthal, C., 1954. Recombination in phage T2: its relationship to heterozygosis and growth. Genetics, 39: 169-184.

10. Trautner, T. A., 1958. Untersuchungen am Heterozygoten des Phagen T1. Z. Vererbungslehre 89: 264-271.

11. Doermann, A. H., and L. Boehner, 1961. An experimental analysis of bacteriophage heterozygotes. (Unpublished manuscript.)

12. Edgar, R. S., 1961. High negative interference and heterozygosis: a study of the mechanism of recombination in bacteriophage T4. Virology, 13: 1-12.

13. Benzer, S., 1955. Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci. U.S., 41: 334-354.

14. Doermann, A. H., and M. Hill, 1953. Genetic structure of bacteriophage T4 as described by recombination studies of factors influencing plaque morphology. Genetics, 38: 79-90.

15. Wollman, E. L., 1953. Sur le déterminisme génétique de la lysogénie. Ann. inst. Pasteur, 84: 281-293.

16. Fredericq, P., 1956. Recherches sur le transfert génétique des propriétés lysogènes par croisement des souches Escherichia coli K.12 et B. C. R. Soc. Biol., 150: 2035-2039.

17. Feynman, R. P., 1960. Personal communication.

18. Adams, M. H., 1959. Bacteriophages, pp. 450-451. Interscience, New York.

19. Delbrück, M., 1940. The growth of bacteriophage and lysis of the host. J. Gen. Physiol., 23: 643-660.

20. Benzer, S. and F. Jacob, 1953. Étude du développement du bactériophage au mayen d'irradiations par la lumière ultra-violette. Ann. inst. Pasteur, 84: 186-205.

21. Edgar, R. S., and R. H. Epstein, 1961. The inactivation by ultraviolet light of an acriflavine sensitive gene function in phage T4D. (Unpublished manuscript.)

22. Edgar, R. S., 1958. Phenotypic properties of heterozygotes in bacteriophage T4. Genetics, 43: 235-248.

23. Beale, G. H., 1954. The genetics of Paramecium aurelia. Cambridge University Press.

FIGURES

FIGURE 1. The Levinthal (9) model for heterozygosis and recombination for distant markers (after Edgar (12)).

FIGURE 2. Edgar's (12) generalized scheme for the production of recombinants for distant (Route I), intermediate (Route II) and close (Route III) markers (after Edgar (12)).

FIGURE 3. A cross involving three closely linked rII markers and the two loosely linked outside markers. The segregation patterns for the outside markers predicted by the rules of classical genetics and by the Edgar scheme are shown.

FIGURE 4. Map showing approximate linkage relationships of the markers used in the crosses described in the text. Map distances not in parentheses are frequencies of recombination in two-factor crosses at normal lysis. Map distances in parentheses are similar measures but are based upon crosses to adjacent markers.

Figure 1

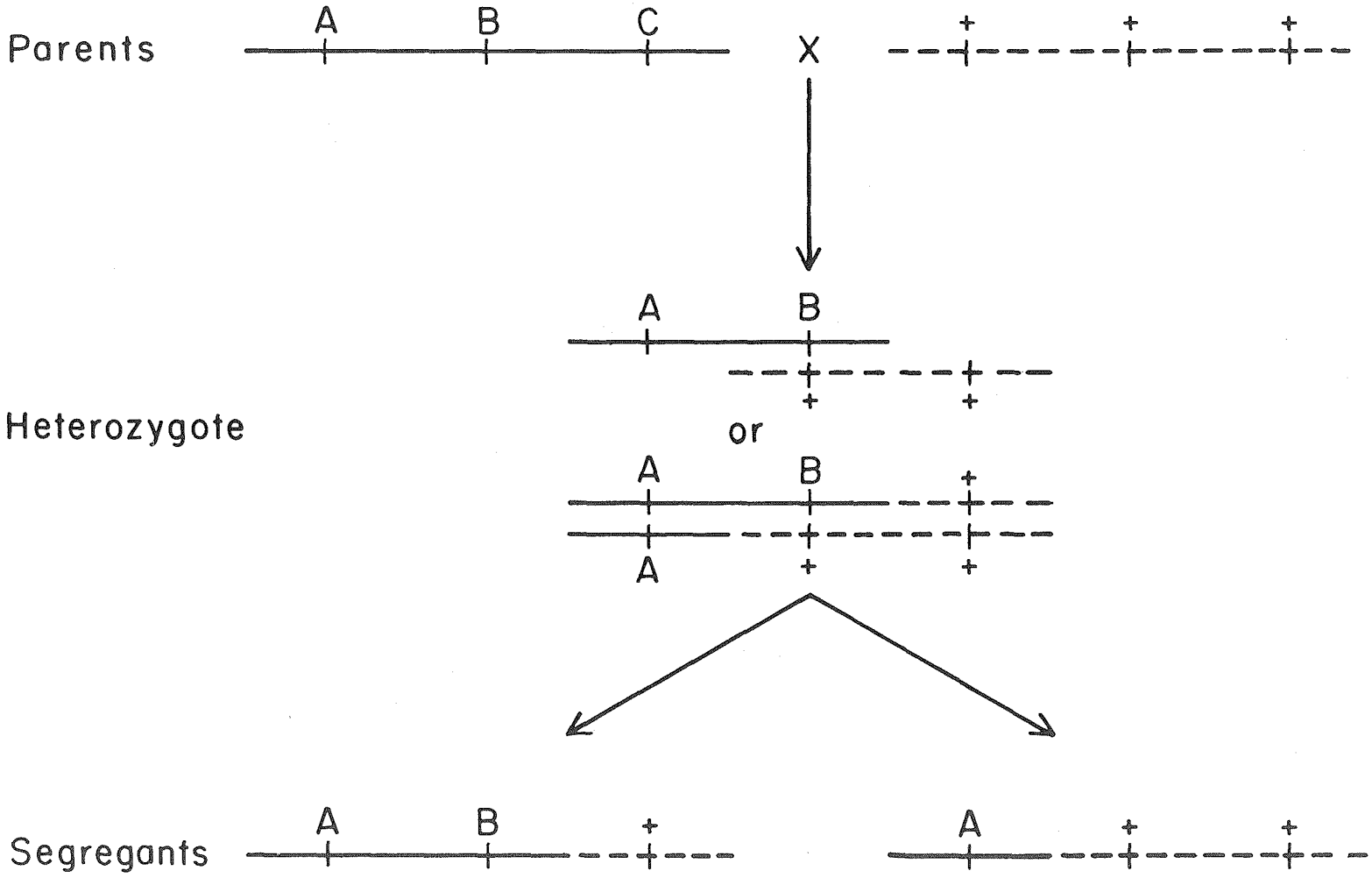
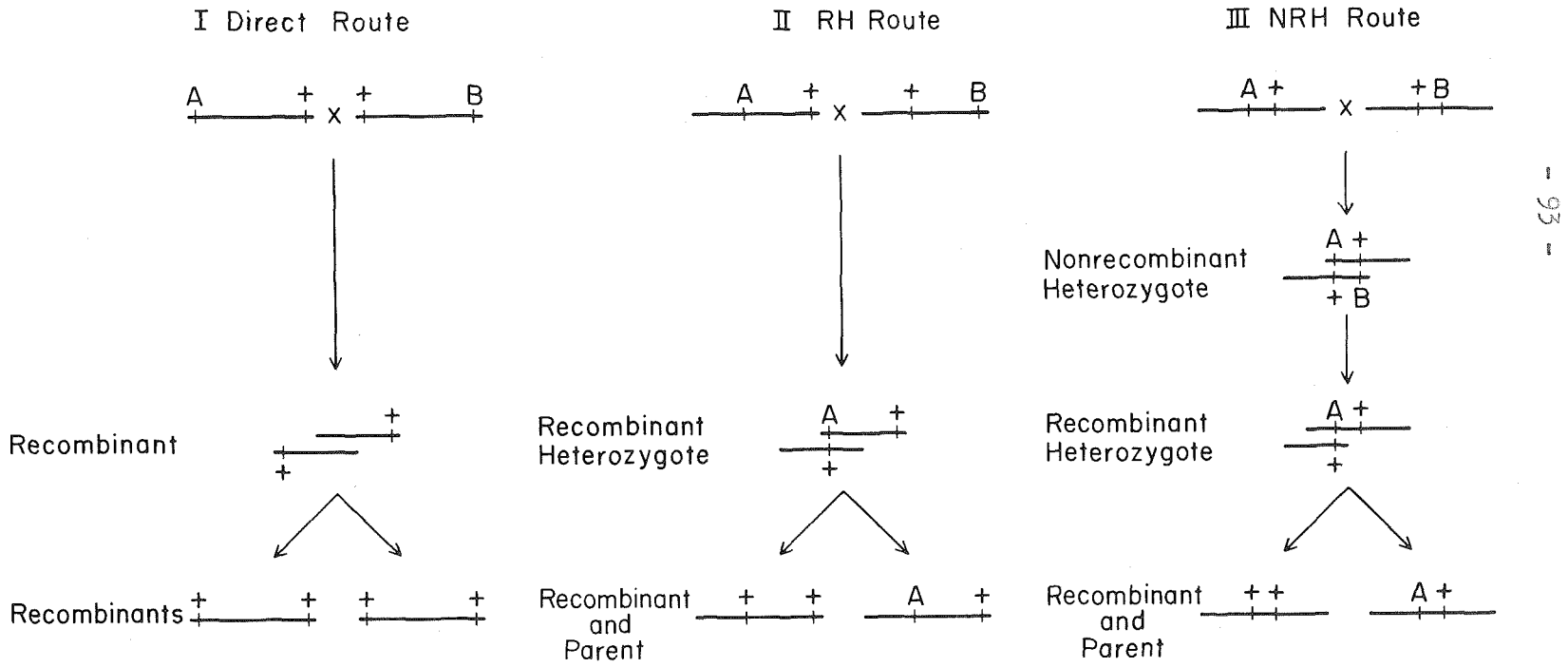


Figure 2


I Direct Route
A + x + B
Recombinant
Recombinants
II RH Route
A + x + B
Recombinant Heterozygote
A +
Recombinant and Parent
+ +
A +
III NRH Route
A + x + B
Nonrecombinant Heterozygote
A +
+ B
Recombinant Heterozygote
A +
Recombinant and Parent
+ +
A +

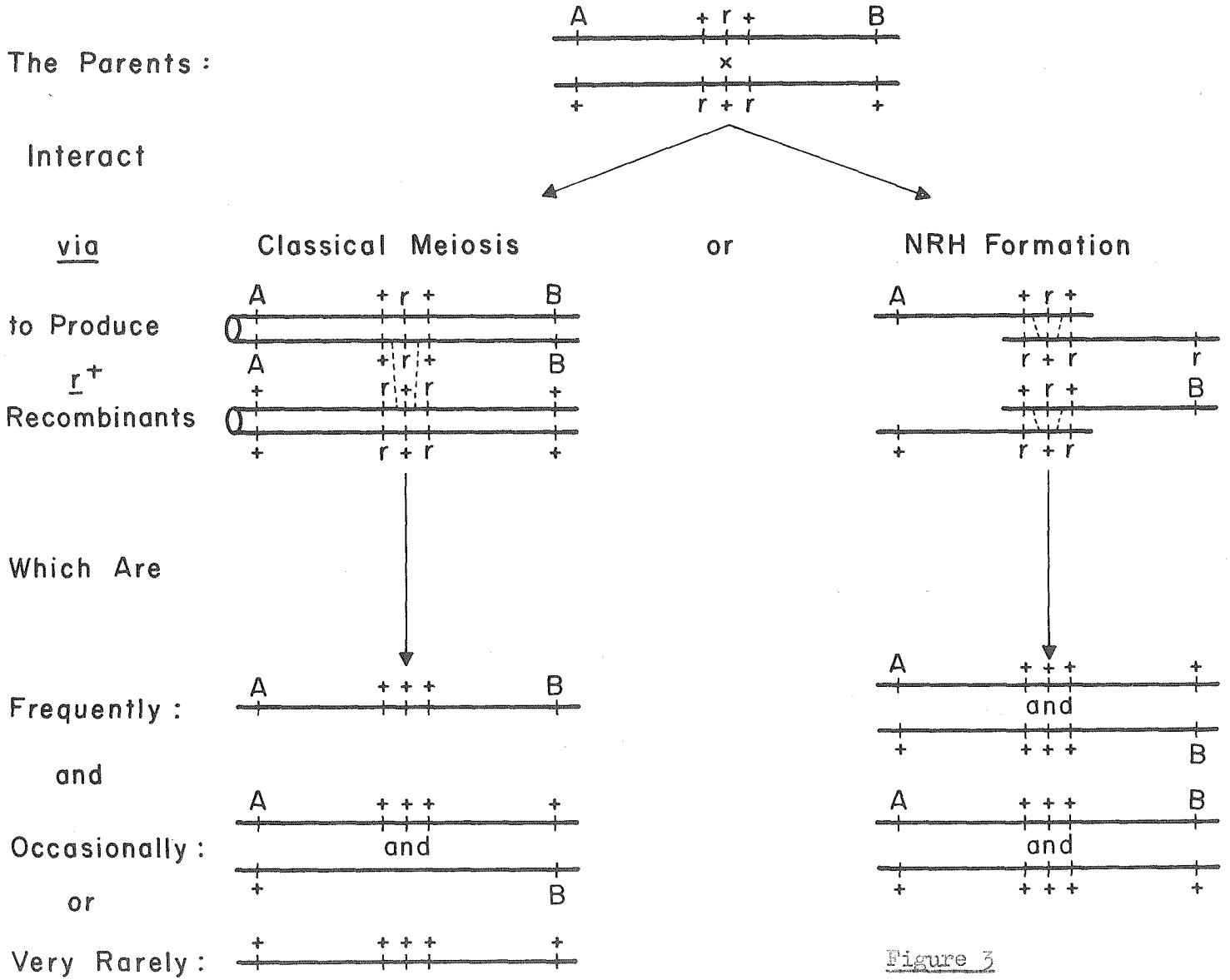

The Parents :
A + r + B
×
+ r + r +
Interact
via
Classical Meiosis
or
NRH Formation
to Produce
r+
Recombinants
A + r + B
A + r + B
+ r + r +
+ r + r
A + r + B
+ r + r
r
+ r + B
+ r + r
Which Are
Frequently :
A + + + B
and
Occasionally :
A + + + +
and
+ B
or
Very Rarely :
+ + + + +
A + + + +
and
+ + + + B
A + + + B
and
+ + + + +

Figure 3

Figure 4