Genetic Studies of Recombining DNA in Pneumococcal Transformation

HARRIETT EPHRUSSI-TAYLOR and THOMAS C. GRAY

From the Developmental Biology Center, Western Reserve University, Cleveland

ABSTRACT   The results of genetic fine structure experiments, performed on the *amiA* locus of *Pneumococcus* are summarized. The peculiar feature of transformation genetics is that a given donor marker mutation transforms with an efficiency characteristic of the mutated site. In spite of this difficulty, mapping procedures have been devised and quantitative recombination studies performed. It is concluded from these studies that transformation, in this locus, is the consequence of frequent, and essentially random exchanges occurring between donor DNA and the chromosomal DNA of the recipient cell. The average length of uninterrupted donor DNA polynucleotide strand which could be inserted into the chromosome of a transformed cell is estimated, from genetic data, to be probably not greater than $3 \cdot 10^5$ daltons (for a double-stranded insertion). It is proposed, on the basis of genetic evidence, that following essentially random exchanges between donor DNA and recipient chromosome, a revision process, specific for certain types of mutated sites, occurs. The revision process appears to remove preferentially donor DNA sequences from the primary recombinant structure, and allow repair along the chromosomal template, leading to low efficiency in the genetic integration of these sites. A mechanism for this "destruction-choice" process is presented, and evidence in support of this mechanism discussed.

With the discovery, published in 1951, of recombination between mutant type III capsular transforming factors, known to be composed essentially if not exclusively of DNA, it was postulated for the first time that recombination at the molecular level involves the reassortment of subunits smaller than the DNA molecules which were effecting these transformations. To the extent that transforming agents could be identified with genes, recombination could be seen to reassort units smaller, also, than genes. At that time, the possibility was raised that recombination of transforming DNA might involve the actual physical exchange of subunits of DNA, much as chromosomal crossing-over is believed to reassort subunits of chromosomes, through breakage and rejoining. An alternate possibility was, however, proposed, namely:

". . . it should nevertheless be kept in mind that no one has the slightest

notion of how even the smallest autoreproducing unit is duplicated, and it is quite possible that the 'crossing-over of molecules' is only apparent. For example, if two transforming agents were properly juxtaposed in the cell, parts of each might be copied for reproduction of a daughter cell." (1) This second possibility was reminiscent of a hypothesis advanced by Belling in 1933 (2) in order to explain crossing-over in higher organisms without invoking chromosomal breaks, and is known as the "copy-choice" theory.

Although today we have rather precise ideas of how DNA and RNA molecules must be reproduced, owing to the flood of experimentation to which the Watson-Crick model of DNA structure gave rise, we are far from having the same degree of precision in our picture of the recombination process. It became clear in 1961 from the work of Meselson and Weigle (3), and Kellenberger, Zichichi, and Weigle (4), that recombination can involve actual physical exchanges between DNA molecules. Further, the work of Benzer (5) showed that the ultimate subunit of recombination is in all probability a nucleotide, or a nucleotide pair. We do not know, however, whether or not copy-choice is also a very important recombination mechanism, and, indeed, there is real need of this general type of mechanism to explain features of genetic data. The physicochemical approach, developed by Meselson and Weigle, and exploited by Bodmer and Ganesan (6), and Fox and Allen (7), and others (8), is hardly suited to study of this question. Yet until it is studied successfully, our knowledge of the recombination process will remain incomplete.

I shall present today the results of the genetic experiments performed by my group, because this is the special contribution I can make to the present symposium. I share my colleagues' enthusiasm for modern experimental techniques, but also am convinced that if we are to use bacterial transformation systems for analyzing molecular recombination, we must first of all have a very detailed genetic description of recombination in transformation. Consequently, one of the major objectives in my laboratory for the past 5 years has been to obtain that description, which has been conspicuously lacking until now. Following presentation of our results, discussion will focus on two questions. (a) Does the image of recombination obtained through genetic crossing coincide with that presented to us on the basis of physical-chemical studies? (b) Is there any evidence for or against some form of copy-choice operating in determining the final structure of a recombinant chromosome?

Good genetic fine structure work requires a suitable experimental system. At the outset, our criterion of a good system was that it should be possible to score quantitatively transformation (or mutations) of wild-type cells to mutant ones, and the reverse: mutant cells back to wild type. Sicard (9) developed a synthetic medium for culture of our pneumococcal strains, with which he was able to ascertain that a class of aminopterin-resistant muta-

tions, long suspected as possessing increased nutritional requirements, were not in fact characterized by increased auxotrophy, but, rather, an extreme sensitivity to an imbalance in the molar ratios of isoleucine, leucine, and valine. The two properties of this class of mutations, i.e. aminopterin resistance and sensitivity to an imbalance of branched amino acids, cannot be separated by recombination. They are the consequence of single mutational events in a particular region of the pneumococcal genome. Owing to this fact, circumstances are exceedingly favorable for genetic fine structure work in the aminopterin resistance region (*amiA* region).

DEFINE EFFICIENCY OF GENETIC INTEGRATION

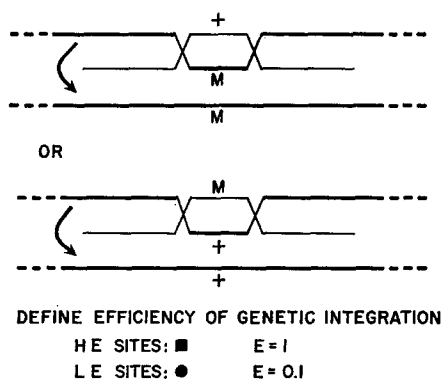H E SITES: ■   E = I
L E SITES: ●   E = 0.I

FIGURE 1. Diagram of reciprocal one-point transformation crosses. HE, high efficiency. LE, low efficiency. Lines represent either the bacterial chromosome, or double-stranded donor DNA molecules.

The One-Point Reciprocal Cross

This type of genetic cross is uniquely possible in genetic systems of unequal parental contribution. We examine the frequency with which a particular mutation, located on a fragment of a genome (such as a DNA molecule), becomes inserted into a whole chromosome. This is what the very first transformation experiments involved. In the crossing system we have developed for the *amiA* region of the pneumococcal genome, we may perform these crosses in two ways: (*a*) normal (wild-type) aminopterin-sensitive cells may be transformed for resistance, with DNA from a resistant donor, and the transformants scored by plating in complete medium containing aminopterin (0.5 to 1 × 10^{-5} M); (*b*) resistant cells may be transformed back to wild type by DNA from aminopterin-sensitive cells, and transformants selectively scored in synthetic medium containing an excess of isoleucine (or leucine or valine). We can measure, in this type of cross, the efficiency with which any marked point is recombined into a complete chromosome, and in our system do this in both possible directions of crossing (Fig. 1). Notice that an odd number of exchanges is required on either side of the marked point, yielding an even number of total exchanges. A feature of these crosses is that there are no apparent restrictions as to how close to the marked site the exchanges need occur. In this, and in all subsequent crosses, variations in the

amount of DNA taken up by cells of a competent culture are corrected for by relating the frequency of the event under study to the frequency of transformation for a standard marker gene located on a different DNA fragment (10, 11). The reference gene used is a streptomycin resistance gene, *str-r41*, and efficiency is expressed as the ratio of *amiA-r* (or *amiA-s*) to *str-r41* transformants. The some 85 *amiA-r* mutations thus far examined prove to fall into two nonoverlapping efficiency classes, whose means differ by a factor of 10: high-efficiency (HE) mutations yielding a ratio of 1 relative to the reference gene, and low-efficiency (LE) mutations, a ratio of 0.1. This result is independent of the direction of the cross. We have already published an extensive report on the distinguishing properties of the two efficiency classes (12), which I shall not review. Suffice it to say that genetic evidence shows that HE mutations may be of two sorts: point mutations and small deletions. LE mutations seem to be, thus far, only point mutations. Efficiency is site-specific, and remarkably constant under a variety of conditions.

Recently, Lacks (13) has reported extensive fine structure work in the maltase locus of *Pneumococcus*, his results differing somewhat from ours. Many of his mutations are deletions, and the proven deletions show efficiencies which range from high to low. Indeed, an inverse correlation is observed between efficiency and the length deleted,[1] small deletions having the efficiency of our HE class (in agreement with our results), and large ones having efficiencies as low as our LE class. Among the mutations believed on the basis of genetic evidence to be point mutations, Lacks discerns 4 efficiency classes, and advances the interesting speculation that these classes correspond to classes of transitions and transversions. The difference in the number of efficiency classes observed in Lacks's experiments, as opposed to ours, may be real, or may be due to difficulties in screening out deletion-type mutations in his experiments. Thus far, in spite of extensive mapping, only one proven deletion has been found in the *amiA* region, suggesting that point mutations in this region very readily result in an altered gene product. This in turn may eliminate a classification problem for us, deletion mutations being virtually absent.

In 1959 (14, 15), an LE mutation in an entirely different region of the pneumococcal genome (optochin resistance *opt-r*, then symbolized by Q^b) was shown to penetrate into the competent cell as readily as the HE *str-r41* gene. This could be done by allowing competent wild-type cells to absorb for 4 min DNA bearing the two genes, destroying unabsorbed DNA with DNAase, washing the cells free of fragments, and placing them in medium at 37 °C for further growth. Samples of the cells were removed at intervals,

---

[1] However, it should be noted that M. B. Rotheim and A. W. Ravin (*Proc. Nat. Acad. Sc.* 1964, **52**, 30) do not find such a correlation. A large multisite mutation in the *str-r* locus shows the same efficiency as the recombining sites which it apparently covers.

lysed, and their marker-gene content assayed on wild-type cells. The *opt-r/str-r41* transformant ratio obtained with the input DNA was 1/8. In the earliest extracted DNA sample, *opt-r* and *str-r41* transformants could be obtained in the assay strain at a ratio of 1/16. Three min later, extracts yielded the two types of transformants at a ratio of 1/50, following which the ratio remained at a value of approximately 1/64. Clearly, the LE gene was being destroyed inside the cell to an extent equivalent to its probability of genetic integration. It was suggested at that time that following penetration into the cell, the LE marker was rapidly destroyed, probably by the recombination process itself. To explain why one marker gene was destroyed during recombination and the other not, it was suggested that the LE gene *opt-r* was a complex locus, of appreciable dimensions, capable of being split by recombination. Indeed, *opt-r* transformants are occasionally observed having a lower level of resistance than that possessed by the DNA donor, indicating that such segregation may be occurring. Were the recombination process to involve breakage of the donor DNA template, followed by insertion in a chromosome and rejoining of broken ends, we could well understand why the intracellular concentration of *opt-r* genes would decrease during recombination. We have every reason to believe that the LE mutations in the *amiA-r* region behave as does the *opt-r* gene, once inside the cell. These mutations are, however, almost certainly point mutations, and we cannot postulate the same mechanism for their destruction. Some peculiar feature of their structure must single them out for destruction, presumably as part of the recombination process. We shall return to this point at the end of my report.

Two-Point Transformation Crosses in the amiA Region

In these crosses we are treating cells bearing one *amiA-r* mutation by DNA from cells bearing an independent *amiA-r* mutation, and examining whether wild-type recombinants are produced (Fig. 2). If the two mutations are at different sites, wild-type recombinants will occur, and their frequency, relative to transformants for the *str-r41* gene, may be scored. To obtain such recombinants an odd number of exchanges must occur on each side of the mutation in the recipient chromosome, as in one-point crosses. However, there is a significant difference between one- and two-point crosses: the presence of a mutation in the donor DNA imposes a restriction on one side, as to where the exchanges may occur. The mutant site in the donor molecule must be excluded. Clearly, the closer the mutant site in the donor is to that in the recipient, the harder it ought to be to exclude the donor site. This is the principle applied for genetic mapping, which can be done quite successfully (16).

One remarkable feature of recombination in transformation emerges from two-point crosses, namely, the role of the efficiency of mutant sites as trans-

forming agents, in determining the actual recombination frequency observed (11, 16). When the two mutations involved in such crosses are of different efficiency, the recombination frequency observed will be dependent on *the efficiency with which the mutant site in the recipient chromosome* can be eliminated from the chromosome. In our experiments, in which we have two efficiency classes differing by a factor of ten, crosses in reverse polarity give wild-type recombinants at frequencies which differ in general by a factor of ten (see Fig. 2). This is a very peculiar result, in view of the fact that the distance between the two mutant sites is really the same in both directions of crossing. In spite of this feature of transformation genetics, it has been possible to map our mutants with complete certainty, by methods which I shall not describe

TWO-POINT CROSSES: SELECTION FOR + +.
ROLE OF EFFICIENCY OF MUTANT SITE RECIPIENT

CROSS A YIELDS >> + + RECOMBINANTS THAN CROSS B,
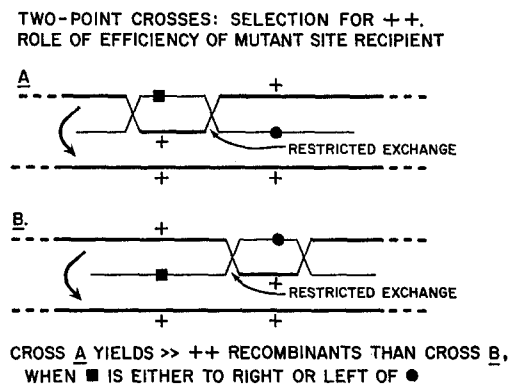WHEN ■ IS EITHER TO RIGHT OR LEFT OF ●

FIGURE 2.   Diagram of reciprocal two-point transformation crosses.

here since they have been published (16). Fig. 3 shows the map obtained with 56 mutations, mapping in 43 sites. You can see that the high-efficiency mutants, the minority class indicated in heavy numerals to the left of the line representing the chromosome, are distributed along the length of the mapped region.

I have stressed above the magnitude of the effect on recombination of the efficiency of the event which eliminates the mutant site in the recipient cell. We can now ask whether the efficiency of the mutant site in the donor DNA plays any role in determining wild-type recombinant frequency in two-point crosses. One type of experiment relevant to this question is shown in Fig. 4. Here we compare the frequency of + + recombinants observed in crosses of the same recipient, bearing either an HE or LE site, (*a*) with a donor DNA bearing an HE site and (*b*) with a donor DNA bearing an LE site showing very low recombination values with the donor site of cross (*a*). In both of these crosses, it is the length *L* which is (*a*) solely or (*b*) primarily determining recombinant frequency. It is found that such crosses yield recombinants at the same frequency, thus showing that the ease with which the donor mutant site is recombined into (or out of) the chromosome plays no role in determining the ultimate frequency of wild-type recombinants.

The fact that the *efficiency of the mutant site in the donor has no role in determining the recovery of + + recombinants* in two-point crosses has constituted a major obstacle in formulating an explanation of what efficiency is about. Obvious hypotheses such as faulty pairing of the donor and recipient strands in the region of LE mutations are excluded by this fact, since faulty pairing would create difficulties whether the LE gene were in the recipient or in the donor. We shall return to this question in a moment.
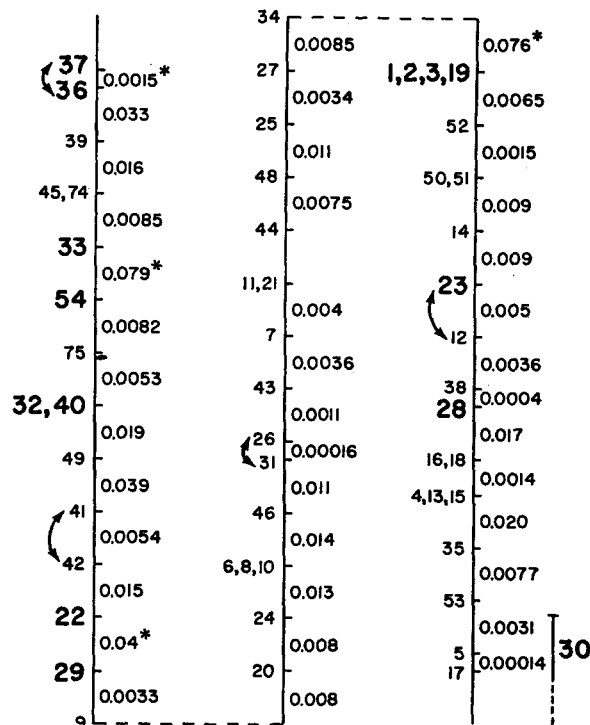

FIGURE 3.   Map of the *amiA* region, made by two-point crossing experiments. *Figure reprinted by permission from Genetics, 1965, 52, 1207.*

Three-Point Transformation Crosses in the amiA region

Let us first of all consider the results of three-point transformation crosses. In these crosses, one of the two parents bears two mutations in the *amiA* region and the other, one, the position of the latter being varied at the experimenter's choice. In practice, it is convenient to use the doubly mutant cells as recipients, crossing them with a series of donor DNA's, bearing, each, a mapped mutation (Fig. 5). The purpose of such experiments is to measure how intensely exchanges are occurring along the heterozygous structure, and whether they are occurring independently of each other.

We can measure the independence and intensity of exchanges in these

TWO-POINT CROSSES. ROLE OF EFFICIENCY OF
MUTANT SITE IN DONOR

THIS CROSS:

COMPARED TO:

FREQUENCIES OF ++ RECOMBINANTS EQUAL.
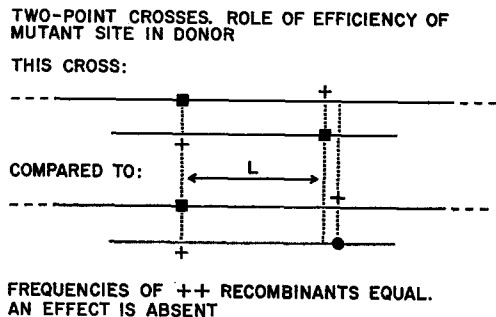AN EFFECT IS ABSENT

FIGURE 4. Diagram of transformation crosses designed to test whether the efficiency of the mutant site in the donor DNA influences the recovery of wild type recombinants in two-point crosses.

crosses in several ways. For example, given the restricted exchange in the first or third three-point cross of Fig. 5, we can calculate what the probability is that a second exchange will occur in the "forbidden" region lying between the mutant sites in the recipient, and determine whether it is the same as the normal recombination frequency of the two genes which delimit the forbidden region, as measured in two-point crosses (Fig. 6). Another calculation we

THREE-POINT CROSSING. SELECTION FOR WILD TYPE

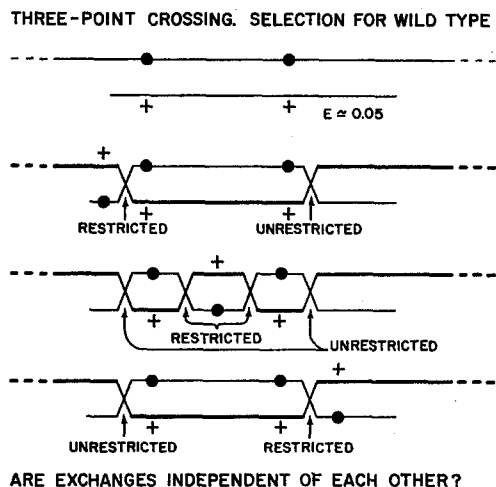ARE EXCHANGES INDEPENDENT OF EACH OTHER?

FIGURE 5. Diagram of three-point transformation crosses. These have been performed also with doubly mutant recipient cells in which one of the mutant sites is a high-efficiency site.

can make is whether those recombinants in which there is no exchange in the forbidden region, in the first or third three-point cross, give rise to exchanges in the restricted region as often as if the recipient contained only the mutated site defining the restricted exchanges (two-point cross). This second comparison is shown in Table I, and one sees that the frequency of the restricted exchange is the same in the two- and three-point crosses.

Another example is furnished by the second three-point cross of Fig. 5. We can measure in two two-point crosses the probabilty of each restricted exchange occurring individually. We can then determine from the three-point cross the frequency with which *both* exchanges occur and see whether it is

equal to the product of the probabilities of each exchange occurring individ-
ually, as expected if exchanges are taking place independently. Our meas-
urements on this type of cross show strong positive interference (one exchange
decreasing the apparent occurrence of a second) when the donor mutant
site lies fairly distant from either site in the recipient. Conversely, if the donor
site is very near one of the recipient sites, interference is absent, or much
lower (Table II). We suspect that in these crosses we are not truly encounter-
ing interference, in the genetical sense of this word, so much as a further
complication arising from the phenomenon of efficiency. Interference would
be expected to be stronger for donor sites *nearest* one of the flanking markers
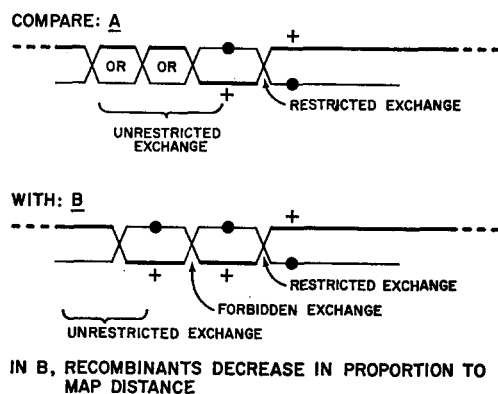of the recipient, not the reverse. It so happens that all the donor sites in our

FIGURE 6. Diagram of one type
of comparison which can be made
of the exchange frequencies in two-
and three-point crosses.

best analyzed three-point crosses of the second type are low-efficiency sites,
and we suspect that had they been HE sites, this anomalous result would not
be found. Crosses to test this point are under way.

In conclusion, therefore, we can say that exchanges are occurring frequently
and essentially independently of each other in the mapped segment we have
studied, the apparent exception being when a cross involves a low-efficiency
site lying midway between two flanking sites in the recipient.

One would like to know what lengths of polynucleotide chains are actually
delimited by the sets of genetic markers used in these crosses. This can only
be a matter of conjecture at present, but since conjectures often reveal which
experiments need doing, it is worthwhile to see what sorts of limits one can
set on the lengths involved. The length of the map which is delimited by
sites *r31* and *r2*, the sites in the recipient used in the crosses examined in
Tables I and II, represents 200 to 300 times the length defined by the lowest
recombination values observed. If one supposes that this latter value corre-
sponds to the frequency of recombination between adjacent base pairs, then
the delimited region would be 200 to 300 base pairs long. Our data show
that it is not unusual for several exchanges to occur within this length.

Turning now to the lowest recombination frequencies measured, which, as we have said, may be between adjacent base pairs, we can derive another approximation of the distances separating two exchanges. The closest pair of low-efficiency markers recombine with a frequency of 1.5×10^{-4}. On the

TABLE I

CROSS OF RECIPIENT r31-r2 BY VARIOUS DONOR DNA'S BEARING
A THIRD amiA-r MUTATION, LYING EITHER TO LEFT OR
RIGHT OF THE REGION DELIMITED BY THE TWO
MUTATIONS IN THE RECIPIENT

| | | Exchanges in restricted region | |
	Donor marker	2-point cross	3-point cross
		%	%
Left of interval r31-r2	r29	37.2	34.3
	r20	29.8	27.1
	r24	26.4	20.7
	r6	15	11.6
Right of interval r31-r2	r52	4.32	2.12
	r50	5.37	2.25
	r12	8.78	11.7
	r28	36.3	28.5

TABLE II

CROSS OF RECIPIENT r31-r2 BY DONOR DNA'S BEARING A
THIRD amiA MUTATION LYING BETWEEN r31 AND r2

| | | Frequencies of 2 restricted switches | | |
| | | Calculated from | Observed in | |
	Donor marker	2-point crosses	3-point cross	O/C
Near r31	r26	0.00228	0.00272	1.19
↑	r43	0.0825	0.0039	0.047
	r7	0.0943	0.00284	0.030
	r11	0.0838	0.00295	0.035
	r25	0.0648	0.00212	0.033
↓	r27	0.078	0.0032	0.041
Near r2	r34	0.023	0.00201	0.087

other hand, a donor DNA molecule contains an average of 10^4 base pairs (for a molecular weight of 6×10^6). This minimal recombination value could be taken to mean that a little more than one random exchange per donor molecule is occurring, the probability of its occurring between any two adjacent bases being $1/10^4$. However, the closest adjacent high-efficiency sites recombine with 10 times this frequency. If we use this value, somewhat more than ten random exchanges per donor molecule might be predicted.

We know, however, that a minimum of two exchanges is required in order to produce a viable recombinant chromosome. We arrive, thus, at values of 2 or something of the order of twenty exchanges per donor molecule, depending on which class of sites we are using for the calculation. If an exchange is the equivalent of breakage and eventual rejoining, the inserted donor fragment could be on the average of molecular weight of either $2 \cdot 10^6$ or $3 \cdot 10^5$, if double-stranded, or half of these values, if single-stranded, again depending on which class of sites we use for calculation.

How can we arrive at a conjecture as to which value (if either) is a realistic estimate of the number of exchanges per donor molecule? One way is to study how genetic linkage decreases as a function of distance, in two-point crosses involving, on the one hand, a high-efficiency recipient, and on the other, a low-efficiency recipient. If linkage decreases rapidly for both types of recipients, we can be sure that a high number of exchanges is involved, but if it decreases slowly, a low number will be indicated. We have shown in the past that the relationship between recombination frequency and distance is linear and additive up to about 25% recombination, irrespective of the efficiency of the mutant site in the recipient (16). Our three-point crosses clearly show that exchanges are frequent and essentially random, so that the departure from additivity, as distances increase beyond this value, is certainly due to additional exchanges occurring along the hybrid region. Inspection of our data shows that long before we have traveled even a third of the mapped region, itself only a portion of a molecule, multiple exchanges are occurring. Further, genetic linkage has virtually disappeard when marker genes are separated by two-thirds of the distance mapped (16). Twenty exchanges per donor DNA molecule would more nearly account for these results than 2, and is probably even not enough. Twenty pairs of exchanges would delimit a polynucleotide chain of the order of $3 \cdot 10^5$ molecule weight, if double-stranded, or one-half of this, if single-stranded.

The characteristics of the curve relating map distance and recombination frequency (16) demonstrate another point, namely, that the region of linear rise in recombination values, as a function of distance between sites in two-point crosses, has the same slope irrespective of the efficiency of the mutation in the recipient, though the absolute values differ by a factor of 10. This indicates that the basic mechanism of exchanges is the same, whether it is a high- or low-efficiency site that is to be excluded from the recipient chromosome. For this reason we feel that we are dealing with a basic "cutting out" process through which donor and recipient polynucleotide sequences are reassorted and that this process must be the same in all crosses. We are left, then, with the possibility that in crosses with low-efficiency sites in the recipient, there is a revision process which sets in posterior to the cutting out process, which rejects the wild-type site in the donor DNA, needed to replace the

mutant site in the recipient in order to obtain a wild-type recombinant. Following the discovery of excision enzymes as the agents of dark repair of ultraviolet damage in DNA (17, 18), their intervention in recombination has been shown to be probable (19). Caffeine, known to inhibit dark repair in *E. coli* and *D. pneumoniae* (20), is without effect on the efficiency of the marker genes of the *amiA* region (20, 21). A number of other purines or purine analogues have been tested for an effect on the efficiency of our marker genes,

CROSS: SUPPOSE PAIRING OF SINGLE STRANDS

PRIMARY RECOMBINANT STRUCTURE YIELDING ++

CUTS
FOLLOWED EVENTUALLY BY JOINING OF ENDS,
AND REPLICATION

CORRECTION STEP, SELECTIVE FOR DONOR STRAND,
9 TIMES OUT OF 10

REPAIR THROUGH LOCAL SYNTHESIS AND JOINING

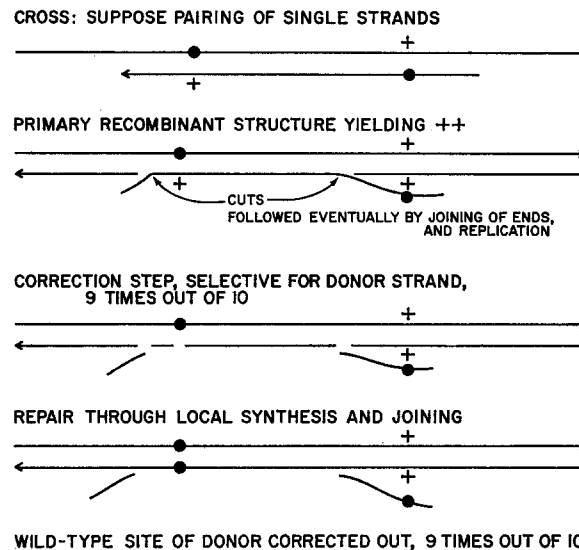WILD-TYPE SITE OF DONOR CORRECTED OUT, 9 TIMES OUT OF 10

FIGURE 7.   Diagram of a recombination mechanism involving pairing of single-stranded donor DNA with the complementary region of one strand of the recipient chromosome-followed by local correcting out of a low-efficiency donor site. The basic exchange mecha, nism creates, in this model, a potential wild-type recombinant. Excision is presumed to occur at the level of the low-efficiency site, removing the donor site preferentially. Repair along the complementary strand derived from the recipient restores the mutant condition of the chromosome. In these, and subsequent diagrams, each line represents a single polynucleotide strand.

and have been found to have none (22). While these experiments lend no support to the idea that dark repair enzymes are involved in determining efficiency, we believe that "correcting out" of low-efficiency donor sites is in fact operating, because such a mechanism can account for all the anomalies of the genetic data thus far encountered. Indeed, we can account for observations cited above, reported in 1960 and 1961 (14, 15), according to which low-efficiency markers undergo destruction inside the cell, during the recombination process. An excision mechanism would account for such destruction very well. A second feature, reported also some years ago, was that a low-efficiency gene is transmitted to progeny earlier than the high-efficiency

gene. We shall see below how this second property can also be accounted for by the excision theory.

Let us now examine the idea of excision being responsible for the efficiency of a particular site as a transforming agent. Let us assume that a primary recombinant structure is formed between single complementary strands of donor DNA and recipient chromosomal DNA, during replication. If the pairing of the complements were in competition with localized synthesis of new DNA, we would arrive at a hybrid structure like that shown in Fig 7.

CROSS: SUPPOSE PAIRING OF SINGLE STRANDS

PRIMARY RECOMBINANT STRUCTURE, NOT YIELDING WILD TYPE

CORRECTION STEP, SELECTIVE FOR DONOR STRAND 9 TIMES
                                                    OUT OF 10
HALF OF CORRECTIONS:

SHOULD CONVERT TO:

WHICH GIVES RISE TO ONE ++ CELL ON REPLICATION.
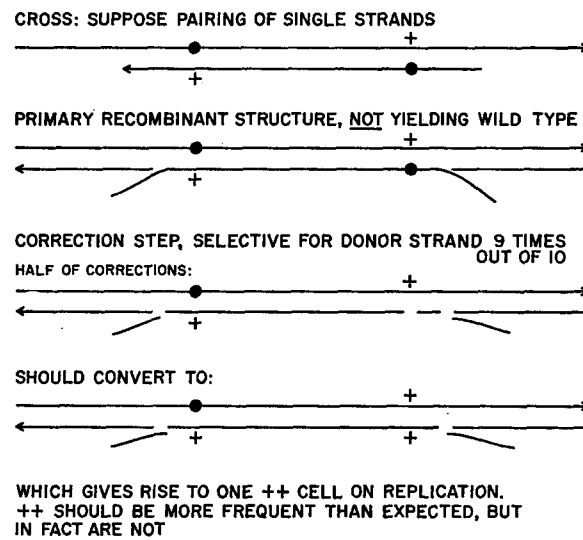++ SHOULD BE MORE FREQUENT THAN EXPECTED, BUT
IN FACT ARE NOT

FIGURE 8. The correcting mechanism postulated in Fig. 7 is expected to convert mutant recombinants to wild-type recombinants. Since this does not happen, highly localized excision cannot be evoked to explain low efficiency.

The unpaired portions of donor DNA could be removed by cuts, or nuclease action, and a hybrid complementary strand formed by joining of ends. At the first cell division posterior to DNA uptake, a pure nontransformed cell would be segregated from a cell bearing a hybrid heteroduplex (one strand unchanged, one strand recombinant). At the second division, another pure untransformed cell would be formed and segregated from a pure transformed one. This model predicts that a donor marker would not be replicated until the second chromosomal division after DNA uptake, and not transmitted to sister cells until third division following DNA uptake, a prediction which has been borne out experimentally for high-efficiency donor genes (15, 23).

Let us now suppose that when a low-efficiency site is included in the primary recombinant structure, in the heteroduplex state, an excision enzyme recognizes some feature of the structure (absence of pairing between one base

of each complement?) and excises one strand in the region recognized, choosing to excise the donor strand nine times out of ten, but choosing the recipient strand for excision, the remaining time. This would account for low efficiency.

The difficulty with this hypothesis as it stands is that it predicts that it should make a difference whether the *donor* site to be excluded in two-point crosses is a high- or low-efficiency mutant. (In fact, the efficiency of the donor mutant site is without consequence for the recombination values observed in two-point crosses.) Fig. 8 shows why the hypothesis makes this prediction. Let us take the case in which the primary recombinant structure has *included*
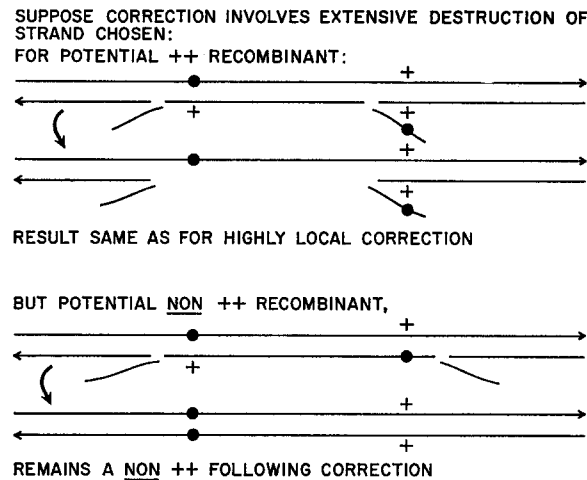

FIGURE 9. The model proposed in Fig. 7 may be modified by supposing that the correcting mechanism excises a long stretch of the donor DNA strand. It then yields the results observed in reciprocal two-point crosses.

the donor-mutated site, and has not prepared the way for emergence of a wild-type recombinant. The proposed correcting mechanism should excise the mutant donor site nine times out of ten, and following repair along the chromosomal complement, cause the formation of more wild-type recombinants than one would expect on the basis of map distance. This is not observed.

This difficulty is removed if we modify our hypothesis to suppose that the excision process is extensive once it gets started, so that an appreciable length of the chosen strand is removed. Fig. 9 shows why. Whether we propose localized or extended excision, the fate of a primary recombinant structure which is a potential wild-type recombinant will be the same: nine times out of ten, the potential wild-type recombinant will be converted back to an untransformed cell. In contrast, a primary recombinant which is not a potential wild type *will not be converted to wild type if the excision process is extensive*, for the excision process will remove the wild-type site of the donor strand as well as

the mutant site, and the repaired chromosome will be like the original recipient chromosome.

Two lines of experimental evidence strongly support the notion of excision as determining efficiency, and show that excision must be extensive.

The first line of evidence bearing on excision comes from the study of the results of two-point crosses in which two mutant sites of different efficiency are in the *cis* position (Fig. 10) as opposed to the *trans* position (Fig. 2). If extensive excision of the donor strand is occurring, we should observed it in the *cis* crosses as a drastic reduction of the probability of the *high-efficiency* site being incorporated into a chromosome. In the cross which we shall consider here, of wild-type cells by *r36-r39* DNA, 70% of all the cells incorporat-

DONOR DNA's:

    r 36   r 39

AND

        r 39   r 33

CROSS:
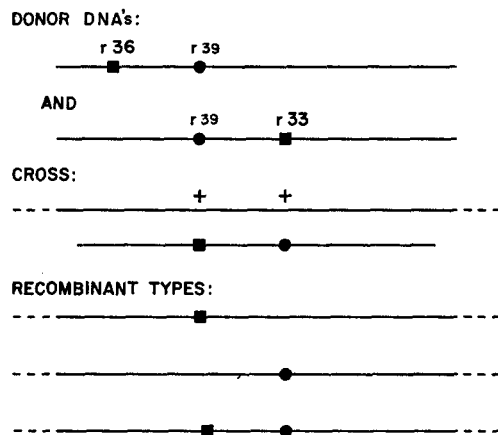
        +   +

RECOMBINANT TYPES:

FIGURE 10. Diagram of two-point crosses with the mutant sites in *cis* position, and where the sites are of different efficiency classes.

ing one site should also incorporate the other, because of their relatively close linkage. If the high-efficiency site retains its efficiency, there should be one transformant bearing this site (*r36*) for each *str-r41* transformant. Of these, 0.3 would be expected not to have included the low-efficiency site (*r39*). Of the remaining 0.7 which ought to contain both, only one-tenth, or 0.07 should in fact also contain site *r39* in view of the low efficiency of the latter. We might expect the remaining 0.63 to bear only the *r36* site as a result of the exclusion mechanism operating for *r39*. Total transformants bearing only the *r39* site would be expected at a ratio of 0.03, relative to *str-r41* transformants. Total *amiA* transformants should be, then, about 1.1 relative to the *str-r41* transformants. The observed total, however, proves to be 0.2. The different types of *amiA* segregants recovered are shown in Table III. It is quite clear that there is a strong deficiency of single transformations for the high-efficiency marker gene, and not for the other classes. The process which excludes the low-efficiency donor site effectively excludes the high-efficiency site when it is in a *cis* relationship to the former site. The extensive excision

mechanism very beautifully accounts for this result, which we have now observed in three other crosses involving different pairs of marker genes in the *cis* position.

One final bit of information should be added concerning the segregants, and that is that once the *r36* mutation has segregated from the *r39* mutation, it shows an efficiency which is entirely normal, namely, 1 relative to the *str-r41* transformants.

The second line of evidence in favor of excision comes from the mode of transmission of acquired genes to daughter cells at successive divisions following DNA uptake. As mentioned above, it was established some time ago that whereas the high-efficiency *str-r41* gene is usually transmitted to daughter

TABLE III

RESULTS OF A TWO-POINT CROSS WITH MUTANT
MARKER GENES IN THE *cis* POSITION

Wild-type cells were transformed by DNA from cells bearing *amiA* mutations *r36* (high-efficiency) and *r39* (low-efficiency), as well as the reference marker *str-r41*.

Recombinant	Frequency relative to to *str-r41*	Expected frequency
r36	0.086	0.3+0.63
r39	0.041	0.03
r36–r39	0.072	0.07
Total	0.199	1.1

cells at the third division following DNA uptake, a low-efficiency gene, *opt-r*, is always transmitted to progeny at the second division (14, 15). This suggests that transformants become homozygous or homoduplex for low-efficiency genes, acquired through transformation, one replication cycle earlier than for high-efficiency genes. The generality of this transmission pattern has now been established for high- and low-efficiency sites in the *amiA* region (24).

This has been done in the following way. Wild-type cells are allowed to fix transforming DNA for 10 min, arresting absorption with DNAase. The cells are then diluted eightfold into fresh medium, and allowed to divide at 37°C. Divisions of these cells are synchronized for at least three fissions following dilution (23). Sonication, applied at appropriate moments, separates cells at successive divisions, producing suspensions which are composed of from 50 to 80% of single cocci, and causing as little as 10% mortality. The moment of transmission of markers can be assessed quantitatively by making cell counts and measuring the numbers of colony-forming units able to give rise to transformants before and after sonication.

Since these (tedious) experiments have been performed quantitatively

with the *str-r41* and *opt-r* marker genes, and their distinctive transmission patterns established, one can use these genes as reference genes for study of the transmission pattern of any other marker gene. Without determining mortality and degree of separation, one can decide whether the transmission pattern is like that of the *str-r41* gene or the *opt-r* gene by measuring the ratios of the various transformants to each other, at the outset, and following soni-
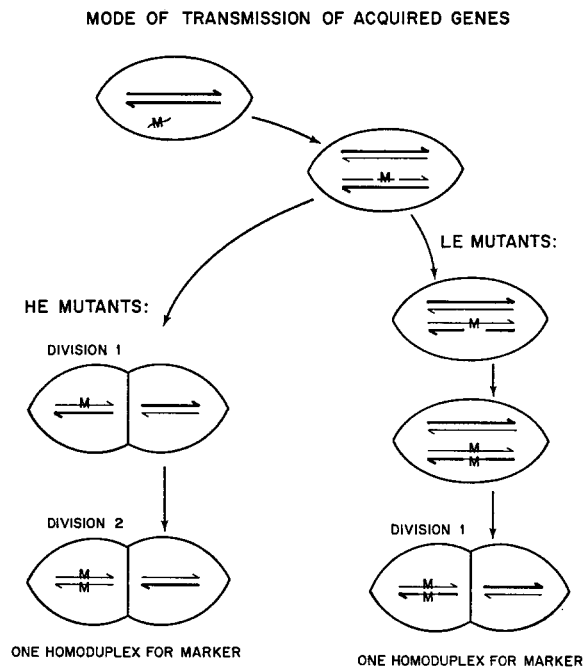
MODE OF TRANSMISSION OF ACQUIRED GENES

FIGURE 11. Diagram showing how the correcting mechanism would lead to early homozygosity if it removed the DNA strand derived from the recipient cell once out of 10 times, and repaired along the template provided by the donor DNA strand.

cation after successive divisions. For example, at the outset, the ratio *opt-r/ str-r41* may be 0.1; the ratio will be unchanged in sonicates made following the first division, but will rise to 0.15 to 0.2 in sonicates made any time following the second division, the amount of rise depending on the effectiveness of sonication in separating the cells. One can then examine whether the proportion of transformants for another genetic marker rises also, in the same manner, or remains constant. The essential task in these experiments is to count enough colonies of each kind of transformant to be sure of the statistical significance of a rise, or the absence of a rise.

In this way it has been established that low-efficiency marker genes of the *amiA* region are transmitted as is the *opt-r* gene, while high-efficiency marker genes are transmitted as is the *str-r41* gene.

This difference in transmission pattern is readily explained if we assume (a) that homologous pairing occurs between single-stranded complements; (b) that high-efficiency genes are rendered homozygous (homoduplex) through semiconservative replication; and (c) that low-efficiency genes are rendered homoduplex through excision and repair before the first cell division takes place (Fig. 11).

DISCUSSION

The picture of genetic recombination at the molecular level, obtained through fine structure mapping techniques, is that of a basic pattern of intense, and essentially random exchanges occurring between the donor DNA fragment, and the recipient chromosome. Calculation of the average length of donor DNA polynucleotide sequence inserted into a recipient chromosome has been set as probably not greater than a double-stranded sequence of molecular weight 3×10^5, or a single-stranded sequence half this molecular weight. However, this calculation is subject to one important assumption, namely, that the minimal recombination value observed in our intensely mapped region corresponds to recombination between adjacent base pairs. We are attempting to verify this assumption experimentally, by several independent means, and hope to have either proof or disproof of its validity before very long. If we accept the calculation as valid, and assume that "insertion" means physical insertion of donor polynucleotide sequence into recipient chromosome, then it would appear that estimation according to which recombination is accompanied by the physical insertion of a single-stranded donor segment of $1 \cdot 10^6$ molecular weight, is too high by nearly an order of magnitude.

Whatever the truth turns out to be in so far as our genetic measurement is concerned, it would seem likely a priori that the density-labeling technique would lead to an overestimate. The experimental conditions in transformation are such that the resolution of the density-labeling technique is decreased, relative to resolution in phage, owing to the admixture of large amounts of unrecombined DNA from recipient cells with the recombinant DNA, creating serious hydrodynamic problems. Further, it would be difficult to ascertain whether the increased density of recombinant molecules recovered in the density-labeling experiments is due to the insertion of a segment of donor polynucleotide sequence in the form of one piece, or due to the insertion of several segments of smaller molecular weight, separated by segments derived from the recipient chromosome, for the CsCl gradient technique is not useful for the study of DNA fragments of molecular weight much lower than 10^6. The experimenter is, thus, going to find those recombinant molecules which his technique enables him to find, and this may prove to be a highly biased sample. While low resolving power is not a limitation of

genetic methods, other kinds of errors can be made with them, as already stressed, so that in conclusion we can only suspect that the inserted donor material is in the form of several rather short polynucleotide sequences, rather than one long insertion. We cannot affirm it.

Turning now to the question of whether copy-choice is occurring, we can concede that this too is still an open question. What our experiments do very strongly indicate is destruction-choice followed by repair, as a fundamental feature of recombination in transformation. It would appear that at the stage of the formation of a primary recombinant structure, i.e. at an early stage in the reaction between donor DNA and recipient chromosome, a mechanism enters into action which destroys certain segments of either donor or recipient polynucleotide strands, showing, however, a strong preference for destroying the donor nucleotide sequence. The signal for this destruction would appear to be the configuration of the region which is heterozygous (heteroduplex) for a particular kind of mutational site. The mutations which evoke the destruction mechanism seem to be point mutations. We would like to suggest that destruction is the effect of error-correcting enzymes, which are recognizing single unpaired bases along a linear hybrid structure, the latter, in all likelihood, a heteroduplex composed of one strand of donor and one of recipient DNA. The correcting process would presumably have to intervene *after the basic mechanism of recombination has occurred*, in order to explain why relative map distances are the same for high- and low-efficiency recipients.

The delayed onset of transmission of acquired genes to progeny, already established some years ago for high-efficiency genes, taken in conjunction with the fact that recombinant DNA molecules appear just a few minutes after DNA uptake, has long suggested that genetic integration in transformation proceeds through the formation of a heteroduplex hybrid. DNA replication would thus be required to obtain a homoduplex transformed chromosome, following which transmission to progeny would occur. The physical data appear to confirm this image of genetic integration (6–8). Our own data on the fate of low-efficiency genes, known to undergo partial destruction in the cell, and to be transmitted one division sooner than are high-efficiency genes, are very beautifully explained by destruction of one of two strands of such a heteroduplex, followed by repair. However, we shall have to obtain physical and chemical characterization of the primary recombinant structure, and physical and chemical proof of excision, in order to be absolutely certain that this is the correct picture.

The extended excision, or destruction-choice mechanism, brings a new unity with respect to genetic data on recombination in transformation, and possibly with respect to other systems as well. In our cross, reported in Table III, with high- and low-efficiency markers in *cis* position we have a stunning

example of what Ravin has called the "marker effect" on recombination (25); i.e., a strong effect on the types of recombinants recovered which is dependent on the spatial arrangement of the markers. Owing to the favorable circumstances of our crossing system, we can see that it is the high-efficiency site which is destroyed to about the same extent as the low-efficiency site, when the two are in the *cis* as opposed to the *trans* position. Thus, marker effects and efficiency can be considered two manifestations of a single phenomenon. The "depressor" region of Green (26) can also be interpreted by this mechanism. In certain recipients, a DNA marked with a *str-r* mutation behaves as a low-efficiency site, while in other recipients as a high. A small genetic difference between donor DNA and recipient, present near the *str-r* gene, could serve as a recognition site for the excision enzyme, rendering transforming efficiency low when titered on a particular strain.

Finally, Whitehouse (27) has developed extensive models of recombination involving heteroduplex DNA, and postulated that gene conversion (in particular polarized gene conversion in Ascomycetes) is due to the action of a correction process on heteroduplex hybrid DNA. Our experiments, in which in addition to genetic crossing we have the unique chance of measuring the specific destruction of certain genetic sites in transforming DNA during the recombination process, give substantive support to these models.

REFERENCES

1. EPHRUSSI-TAYLOR, H., Genetic aspects of transformation of pneumococci, *Cold Spring Harbor Symp. Quant. Biol.*, 1951, **16**,445.
2. BELLING, J., Crossing over and gene rearrangement in flowering plants, *Genetics*, 1933, **18**, 388.
3. MESELSON, M., and WEIGLE, J. J., Chromosome breakage accompanying genetic recombination in bacteriophage, *Proc. Nat. Acad. Sc.*, 1961, **47**,857.
4. KELLENBERGER, G., ZICHICHI, M. L., and WEIGLE, J. J., Exchange of DNA in the recombination of bacteriophage λ, *Proc. Nat. Acad. Sc.* 1961, **47**, 869.
5. BENZER, S., The elementary units of heredity, *in* The Chemical Basis of Heredity, (W. D. McElroy and B. Glass, editors); Baltimore, The Johns Hopkins Press, 1957, 70.
6. BODMER, W. F., and GANESAN, A. T., Biochemical and genetic studies of integration and recombination in *Bacillus subtilis* transformation, *Genetics*, 1964, **50**, 717.
7. FOX, M. S., and ALLEN, M. K., On the mechanism of deoxyribonucleate integration in pneumococcal transformation, *Proc. Nat. Acad. Sc.*, 1964, **52**, 412.
8. NOTANI, N., and GOODGAL, S. H., On the nature of recombinants formed during transformation in *Hemophilus influenzae*, *J. Gen. Physiol.*, 1966, **49**, No. 6, pt. 2, 197.

9. SICARD, A. M., A new synthetic medium for *Diplococcus pneumoniae* and its use for the study of reciprocal transformations at the *amiA* locus, *Genetics*, 1964, **50**, 31.

10. ROTHEIM, M. B., and RAVIN, A. W., The mapping of genetic loci affecting streptomycin resistance in Pneumococcus, *Genetics*, 1961, **46**, 1619.

11. LACKS, S., and HOTCHKISS, R. D., A study of the genetic material determining an enzyme activity in Pneumococcus, *Biochim. et Biophysica Acta*, 1960, **39**, 508.

12. EPHRUSSI-TAYLOR, H., SICARD, A. M., and KAMEN, R., Genetic recombination in DNA-induced transformation of Pneumococcus. I. The problem of relative efficiency of transforming factors, *Genetics*, 1965, **51**, 455.

13. LACKS, S., Integration efficiency and genetic recombination in pneumococcal transformation, *Genetics*, 1966, **53**, 207.

14. EPHRUSSI-TAYLOR, H., On the biological functions of deoxyribonucleic acid, *in* Microbial Genetics, (W. Hayes and R. C. Clowes, editors), Cambridge, The University Press, 1960, 132.

15. EPHRUSSI-TAYLOR, H., Etude de la recombinaison à l'échelle moléculaire dans la transformation bactérienne, *J. Chim. Phys.* 1961, **58**, 1090.

16. SICARD, A. M., and EPHRUSSI-TAYLOR, H., Genetic recombination in DNA-induced transformation of Pneumococcus. II. Mapping the *amiA* region, *Genetics*, 1965, **52**, 1207.

17. SETLOW, R. B., and CARRIER, W. L., The disappearance of thymine dimers from DNA: an error-correcting mechanism, *Proc. Nat. Acad. Sc.*, 1964, **51**, 236.

18. BOYCE, R. P., and HOWARD-FLANDERS, P., Release of ultraviolet light-induced thymine dimers from DNA in *E. coli* K12, *Proc. Nat. Acad. Sc.*, 1964, **51**, 293.

19. CLARK, A. J., and MARGULIES, A. D., Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12, *Proc. Nat. Acad. Sc.*, 1965, **53**, 451.

20. LOUARN, J-M., and SICARD, A. M., Sur la sensibilité différentielle des marqueurs génétiques aux rayonnements ultraviolets, et les mecanismes de réactivation au cours de la transformation chez *Diplococcus pneumoniae*, *Compt. rend. Acad. sc.* 1965, **261**, 852.

21. EPHRUSSI-TAYLOR, H., unpublished data.

22. BAKER, J., unpublished data.

23. EPHRUSSI-TAYLOR, H., The mechanism of deoxyribonucleic acid-induced transformations, *in* Recent Progress in Microbiology, Symposia of the VII International Congress of Microbiology, (G.Tunevall, editor), Stockholm, Almqvist and Wiksells, 1958, 51.

24. EPHRUSSI-TAYLOR, H., Genetic recombination in DNA-induced transformation of Pneumococcus. IV. The pattern of transmission and phenotypic expression of high- and low-efficiency donor sites in the *amiA* locus, *Genetics*, 1966, in press.

25. RAVIN, A. W., and IYER, V. N., Genetic mapping of DNA: influence of the mutated configuration on the frequency of recombination along the length of the molecule, *Genetics*, 1962, **47**, 1369.

26. GREEN, D. M., A host-specific variation affecting relative frequency of transformation of two markers in Pneumococcus, *Exp. Cell Research*, 1959, **18**, 466.

27. WHITEHOUSE, H. L. K., Crossing-over, *Sc. Progr.* 1965, **53**, 285.