

GENETIC RECOMBINATION IN *ESCHERICHIA COLI*

I. RELATION BETWEEN LINKAGE OF UNSELECTED MARKERS AND MAP DISTANCE

C. VERHOEF AND P. G. DE HAAN

Laboratory of Microbiology, The State University, Utrecht (The Netherlands)

(Received September 9th, 1965)

SUMMARY

A relation between linkage frequency of an unselected marker and transfer time based on a physical exchange of genetic material was developed for *Escherichia coli* crosses. Crosses performed under standardised conditions have shown that the relation was valid. The linkage frequency is determined by two constants and one parameter. The constants are the incorporation frequency for donor fragments and the average number of breakage events per unit length. The variable in the relation is the distance in time units between selected and unselected marker.

INTRODUCTION

Linkage data in bacterial crosses are obtained from experiments in which recombinants for a distal Hfr marker are selected and unselected markers located between origin of the Hfr chromosome and selected marker are scored. The frequency of the unselected markers among the recombinants rapidly decreases with increasing distance from the selected marker (compare ref. 5). Proximal markers which are transferred more than 10 min before the selected marker are inherited randomly with a frequency which is in most crosses about 50%. The exact relationship between linkage frequency and physical distance in such crosses is still obscure. In this paper we will present a mathematical relation between distance in time units and linkage frequency based on a physical exchange of genetic material during recombination.

MATERIALS AND METHODS

Strains

The donor strain employed was Hfr R₄ (prototrophic and streptomycin sensitive¹⁵).

The acceptor strains employed were: K12 *arg⁻ ade_k⁻ thr⁻ leu⁻ pro_A⁻*; K12 *ade_k⁻ thr⁻ leu⁻ pro_A⁻*; B *ade_k⁻ thr⁻ leu⁻ pro_B⁻*. Both K12 strains are mutants from strain

P 678 (ref. 8); the B strain was a derivative of strain Bc 258 (ref. 1). All acceptor strains were streptomycin resistant.

The symbols of the markers represent: *ade_k*, adenylosuccinate synthesis; *arg*, arginine; *leu*, leucine; *pro*, proline; *thr*, threonine.

Methods of crossing

Overnight cultures of donor and acceptor strains in nutrient broth were diluted tenfold in prewarmed broth and incubated for 90 min on an inclined turntable. Mating mixtures were obtained by mixing 1 vol. of donor with 10 vol. of acceptor cells. In the K12×B cross the cultures were concentrated tenfold before mixing. Five minutes after mixing, the non-rotated suspension was diluted (500× in K12×K12; 100× in K12×B crosses) into prewarmed broth to prevent further contact formation. A transfer time of 50 min was allowed when *ade_k*⁺ recombinants were selected, against 60 min for *arg*⁺ recombinants.

After the transfer period the mating was interrupted by violently shaking for 60 sec on a microid flask shaker. Streptomycin was then added to a final concentration of 100 µg/ml and zygotes were incubated at 37° for an additional period of 60 min on a turntable to obtain full expression of the biochemical characters in the recombinants. During the last period the mixture was violently shaken at intervals of 10 min to prevent the formation of contacts. The 60-min post-aeration period was also necessary to obtain pure colonies on the selection plates.

Selection of recombinants

Samples of 0.3 ml were plated on prewarmed minimal medium supplemented with streptomycin (100 µg/ml) and the necessary growth factors for the unselected markers.

The optimal concentrations of the growth factors which as far as possible prevented the loss of recombinants due to viability effects were: adenine, 20 µg/ml; DL-leucine, 80 µg/ml; DL-proline, 60 µg/ml; and DL-threonine, 100 µg/ml. Large Petri dishes (diameter 11 cm) with 80 ml of medium were used for the same reason. The experiment were performed in a constant temperature room (37°). Recombinants were scored for unselected markers after 48 h of incubation.

MATHEMATICAL ANALYSIS

Two kinds of mechanisms have been proposed to account for genetic recombination in bacteria and bacteriophage. The first theory is that the recombinants are formed by a *copy-choice* mechanism in which the new DNA is first synthesized by replication of one parental DNA molecule followed by a switch to the second double helix, so that the remainder of the new DNA molecule is synthesized by the replication of the second DNA helix. The copy-choice theory implies that no physical exchange of pre-existing genetic material takes place. There is only indirect evidence in favor of the copy-choice mechanism^{7,11}.

The second mechanism involves the physical exchange of pre-existing genetic material between homologous chromosomes. The recombinants are formed by breakage and reunion of DNA molecules. The evidence for the *breakage and reunion* mechanism comes from studies on recombination in bacterial crosses¹⁶, in bacterial

transformation^{4,19} and in phage^{12,10,6,13}. The information which is available about the mechanism of recombination after conjugation is unfortunately too limited to allow any definite conclusions. The experimental evidence however points into the direction of the breakage and reunion model. We will therefore base our mathematical theory on a model which is based on the physical exchange of genetic material.

For simplicity we will divide the recombination process in the zygote into three hypothetical steps. The first step is the pairing of the donor and acceptor DNA followed by an enzymatic process which breaks donor and acceptor DNA at corresponding points. We will further designate this process as the breakage event and we will assume that the breakage points are formed at random. The breakage events are followed by a process in which the DNA segments are rejoined so that one complete chromosome is again obtained. The segments which are not used for incorporation are destroyed. We will further assume that the breakage events are not always followed by the incorporation of donor segments but in some cases the acceptor segment is used. We will assume a fixed probability (α) for the incorporation of a donor segment. The incorporation frequency of a segment is thus independent of the length of the segment and it is also independent of the incorporation events of adjacent segments.

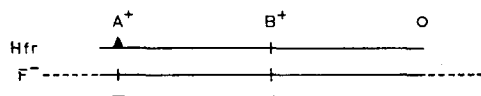


Fig. 1. Diagram of a two-point cross. \circ , origin; \blacktriangle , selected marker.

Let us now consider a hypothetical cross as presented in Fig. 1, in which A^+ recombinants are selected and the marker B is scored as unselected marker. Thus all zygotes which will give rise to A^+ recombinants have received the complete donor fragment from origin to A^+ and the frequency of the unselected donor marker B^+ among the recombinants is thus solely due to recombination events in the zygotes.

We must first consider what happens to the ends of the Hfr fragment. There is experimental evidence that both ends of the donor fragment initiate an obligatory breakage event. Blender experiments have shown that the incorporation frequency of a proximal marker is the same as the incorporation frequency of more distal markers¹⁴. This shows that the distance origin-proximal marker has no influence on the incorporation of that marker and this can only be explained when the probability for the breakage event at the origin is very high. We will assume that the breakage event at the origin is obligatory (*i.e.* the probability = 1). A high probability of cross-over at the distal end of the Hfr chromosome was also observed in blender experiments². We will assume that the distal end also causes an obligatory breakage event.

The two obligatory breakage events at both ends of the Hfr fragment are accompanied by random breakage events occurring between origin and distal end of the Hfr fragment. As the number and the distribution of additional breakage events in the terminal segments have no influence on the incorporation of the markers A^+ and B^+ , we need only to consider the distribution of breakage events in the segment A^+-B^+ .

Let μ be the distance in time units between selected (A^+) and unselected marker (B^+); k be the average number of breakage events per unit length in a population of

zygotes; α be the fixed probability that a donor fragment is incorporated into the recombinant. The average number of breakage events in a segment with length μ is then $k\mu$. The probability for exactly n breakage events in the segment A-B is given by the Poisson distribution:

$$P(n) = \frac{(k\mu)^n}{n!} e^{-k\mu} \quad (1)$$

if the events are randomly distributed over the segment.

Let us now consider a zygote with n breakage events in the segment A-B. An A⁺B⁺ recombinant is obtained from such a zygote when the segments are rejoined so that A⁺ and B⁺ are incorporated into the acceptor chromosome. This requires the simultaneous incorporation of the two segments on which A⁺ and B⁺ are located and the probability for this process is α^2 . The fate of the remainder of the segments is of no importance. The probability to obtain an A⁺B⁺ recombinant from a zygote with n breakage events in the segment A-B is then:

$$P_n(\text{A}^+\text{B}^+) = \alpha^2 \times P(n) = \alpha^2 \frac{(k\mu)^n}{n!} e^{-k\mu}$$

An A⁺B⁺ recombinant may also be obtained from a zygote with no event in the segment A-B. This requires the incorporation of one donor segment only and the probability for this process is:

$$P_0(\text{A}^+\text{B}^+) = \alpha P(0) = \alpha e^{-k\mu}$$

The total probability for an A⁺B⁺ recombinant is:

$$P(\text{A}^+\text{B}^+) = \alpha \times P(0) + \alpha^2 [P(1) + P(2) + P(3) + \dots + P(n)] = \\ \alpha e^{-k\mu} + \alpha^2 (1 - e^{-k\mu}) = \alpha [\alpha + (1 - \alpha)e^{-k\mu}]$$

The probability $P(\text{A}^+)$ for an A⁺ recombinant (B⁺ or B⁻) is α and the frequency (β) of the unselected donor marker among A⁺ recombinants is thus given by Eqn. 2:

$$\beta = \frac{P(\text{A}^+\text{B}^+)}{P(\text{A}^+)} = \alpha + (1 - \alpha)e^{-k\mu} \quad (2)$$

A⁺B⁻ recombinants can only be obtained from zygotes with one or more breakage events in the segment A-B. This requires the incorporation of the donor segment on which A⁺ is located (probability = α) and the incorporation of the acceptor segment on which the B⁻ marker is found (probability = $(1 - \alpha)$).

The probability for an A⁺B⁻ recombinant is then:

$$P(\text{A}^+\text{B}^-) = \alpha(1 - \alpha)(1 - e^{-k\mu})$$

and the linkage frequency of an unselected acceptor marker is:

$$(1 - \beta) = (1 - \alpha)(1 - e^{-k\mu}) \quad (3)$$

Eqn. 3 may also be obtained directly from Eqn. 2.

A plot of linkage frequency (β) against μ is expected to give an exponential curve which asymptotically approaches the value of α when μ becomes very large (compare also Fig. 3). The incorporation frequency α which is assumed to be constant

for different donor segments is equal to the linkage frequency for markers at genetic equilibrium.

Logarithmic transformation of Eqn. 2 gives:

$$^{10}\log(\beta - \alpha) = ^{10}\log(1 - \alpha) - k\mu \ ^{10}\log e \quad (4)$$

According to Eqn. 4, a plot of $^{10}\log(\beta - \alpha)$ against μ is expected to give a straight line intersecting the vertical axis at the point $^{10}\log(1 - \alpha)$.

RESULTS

In all crosses presented in this paper, recombinants were selected for the distal markers *arg* or *ade_k*, whereas proximal markers were scored as unselected markers (Fig. 2). Thus all zygotes which gave rise to the desired recombinants had received the complete donor fragment from origin to selected marker. The frequency of the unselected markers in the recombinants is thus solely due to recombination events in the zygotes.

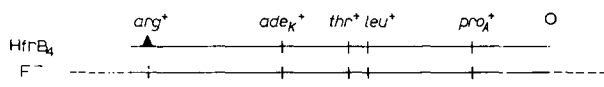


Fig. 2. Diagram of a Hfr $R_4 \times K12$ *arg-ade_k-thr-leu-pro_A-* cross. O, origin; ▲, selected marker.

The distances in time units between the various markers, as given in Figs. 3 and 4 were taken from TAYLOR's map¹⁸. The exact position of the *ade_k* marker is not given on TAYLOR's map and its transfer time was therefore determined in blender experiments. The arginine marker is identical with one of the markers in the *arg A,C,F,H* cluster^{18,17}. Three experiments were performed to test the validity of the

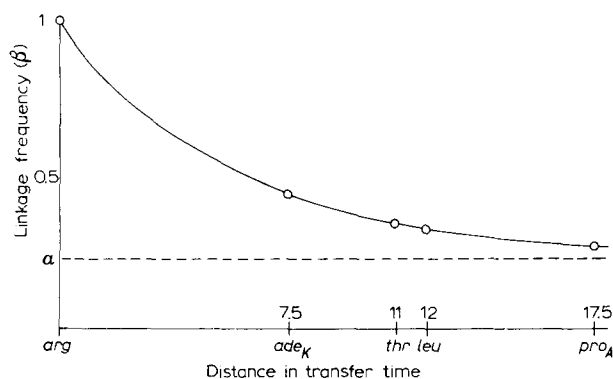


Fig. 3. Relation between linkage frequency and distance in transfer time in a Hfr $R_4 \times arg-ade_k-thr-leu-pro_A-$ cross. The selected donor marker was *arg*⁺.

relation between distance in time units and linkage frequency as given in Eqn. 2. Hfr R_4 was used as donor strain and the K12 strains *arg-ade_k-thr-leu-pro_A-* and *ade_k-thr-leu-pro_A-* and the B strain *ade_k-thr-leu-pro_B-* were used as acceptor strains. The crosses were performed as given under MATERIALS AND METHODS.

The number of recombinants and the linkage frequencies of the unselected

markers is given in Table I. In Fig. 3 the frequencies of the unselected donor markers among *arg*⁺ recombinants from the Hfr R₄ × *arg*[−]*ade*_k[−]*thr*[−]*leu*[−]*pro*_A[−] cross are plotted against the distances in time units. It may be seen from this figure that the frequency of an unselected donor marker rapidly decreases with increasing distance from the selected marker. The curve also strongly suggests that it asymptotically approaches a minimum value (α).

For a plot of $^{10}\log (\beta - \alpha)$ against μ , the exact value of α must be calculated from the data given in Fig. 3 and Table I.

TABLE I
FREQUENCIES OF UNSELECTED DONOR MARKERS WITH Hfr R₄ AS DONOR

| Acceptor strain | Number of recombinants tested | Frequency of donor marker | | | | | |
|---|-------------------------------|---------------------------|-------------------------|------------|------------|-------------------------|-------------------------|
| | | <i>arg</i> | <i>ade</i> _k | <i>thr</i> | <i>leu</i> | <i>pro</i> _A | <i>pro</i> _B |
| K12 <i>arg</i> [−] <i>ade</i> _k [−] <i>thr</i> [−] <i>leu</i> [−] <i>pro</i> _A [−] | 848 | 1 | 0.446 | 0.351 | 0.328 | 0.269 | — |
| K12 <i>ade</i> _k [−] <i>thr</i> [−] <i>leu</i> [−] <i>pro</i> _A [−] | 1512 | — | 1 | 0.731 | 0.688 | 0.562 | — |
| B <i>ade</i> _k [−] <i>thr</i> [−] <i>leu</i> [−] <i>pro</i> _B | 2227 | — | 1 | 0.257 | 0.178 | — | 0.023 |

The value of α was graphically determined by plotting the function $k^{10}\log e = -1/\mu^{10}\log (\beta - \alpha)/(1 - \alpha)$ for each pair of parameters (β and μ) and different values of α .

For each pair of parameters (e.g. $\mu = 7.5$ min, $\beta = 0.446$) a curve was obtained. Theoretically, the four curves should intersect each other at the same point. This point represents the value of α , corresponding to the value for the genetic equilibrium. Due to experimental errors in the determinations of transfer times and linkage frequencies, this is not the case. The intersection points were scattered around the value $\alpha = 0.23$. This value is thus the best approximation for α in this cross.

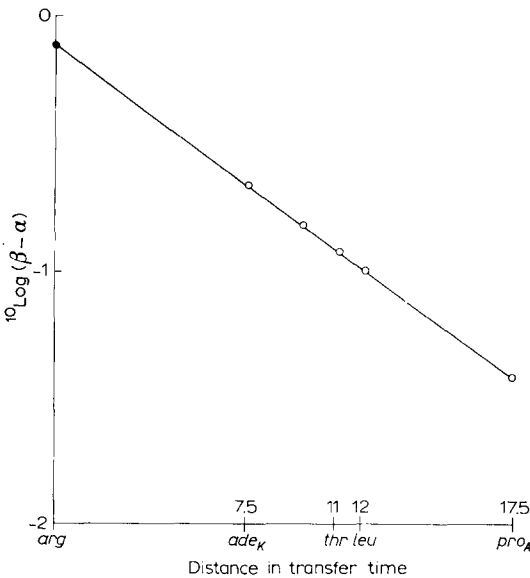


Fig. 4. Regression line of $^{10}\log (\beta - \alpha)$ on distance in transfer time. The same cross as in Fig. 3. ●, the theoretical intersection point ($^{10}\log (1 - \alpha)$) with the vertical axis.

The plot of $^{10}\log(\beta - 0.23)$ against distance in time units is given in Fig. 4. It may be seen from this figure that all points lay on a straight line. The slope (k) of the line was calculated with the least-squares method. The value of k in this experiment was 0.172 breakage events per min transfer time. It may also be seen that the regression line of $^{10}\log(\beta - \alpha)$ on μ intersects the vertical axis very near to the theoretical intersection point $^{10}\log(1 - 0.23)$. With the constants $\alpha = 0.23$ and $k = 0.172$ the linkage frequencies of the four unselected markers were calculated from Eqn. 2. Table II gives the observed and calculated frequencies. The difference between observed and calculated frequencies is less than 2% from the observed frequencies.

TABLE II

OBSERVED AND CALCULATED FREQUENCIES (β) OF UNSELECTED DONOR MARKERS AMONG arg^+ RECOMBINANTS IN A Hfr $R_4 \times arg^-ade_K^-thr^-leu^-pro_A^-$ CROSS

| Marker | Distance in transfer time (min) from <i>arg</i> | β observed | β calculated |
|--------------|---|------------------|--------------------|
| <i>ade_K</i> | 7.5 | 0.446 | 0.443 |
| <i>thr</i> | 11.0 | 0.351 | 0.347 |
| <i>leu</i> | 12.0 | 0.328 | 0.328 |
| <i>pro_A</i> | 17.5 | 0.269 | 0.268 |

848 arg^+ recombinants were scored for unselected markers. The donor frequencies were calculated with the aid of Eqn. 3 with $\alpha = 0.23$ and $k = 0.172$.

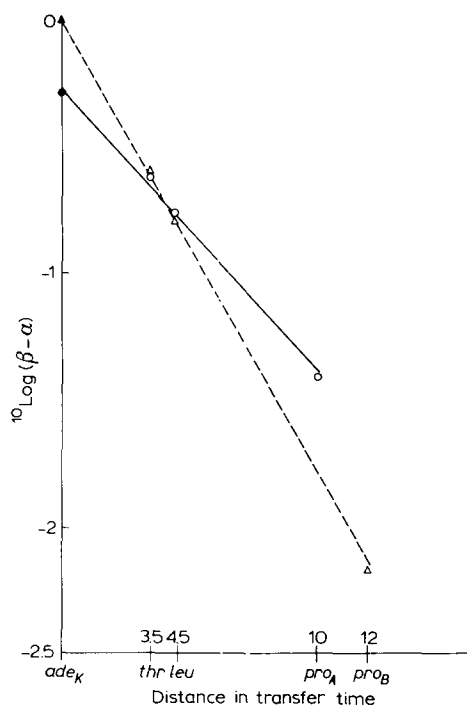


Fig. 5. Regression lines of $^{10}\log(\beta - \alpha)$ on distance in transfer time. \circ — \circ , Hfr $R_4 \times K12 ade_K^-thr^-leu^-pro_A^-$; \triangle — \triangle , Hfr $R_4 \times B ade_K^-thr^-leu^-pro_B^-$; \bullet and \blacktriangle , theoretical intersection points.

In Fig. 5 a plot of $^{10}\log(\beta - \alpha)$ against μ is given for two other crosses with Hfr R_4 as donor. In the first cross, K12 *ade_k-thr-leu-pro_A*⁻ was used as acceptor and 1512 *ade_k*⁺ recombinants were scored for the unselected markers *thr*, *leu* and *pro_A* (compare Table 1). The constants in this cross were: $\alpha = 0.5$ and $k = 0.211$.

In the second cross strain B *ade_k-thr-leu-pro_H*⁻ was used as acceptor and 2227 *ade_k*⁺ recombinants were scored for unselected markers. The constants in this cross were: $\alpha = 0.019$ and $k = 0.378$.

The straight lines given in Fig. 5 show that the relation given in Eqn. 4 is valid for both crosses.

DISCUSSION

The results of the three crosses presented in this paper show that the relation between distance in time units and linkage frequency as given in Eqn. 2 is valid for crosses of the type given in Fig. 2. The relation is also valid for crosses with donors which transfer their chromosomes in the opposite direction as Hfr R_4 . This was confirmed by plotting $^{10}\log(\beta - \alpha)$ against μ with data taken from the literature (*e.g.* ref. 9, p. 227).

The validity of the relation given in Eqn. 3 shows that *any* number of events in the segment A-B may lead to the separation of the markers A and B. The copy-choice model predicts that only the *odd* numbers of events can separate two markers whereas all events of even order would lead to the simultaneous incorporation of the markers A and B. The fact that even numbers of events may separate two markers is thus in contradiction with the copy-choice model of recombination. The results presented in this paper do not, however, prove our model of recombination. Other models, for instance a random fragmentation of the Hfr fragment during or after transfer, and the assumption that each segment is incorporated with a probability α , will also lead to the Eqns. 2 and 3.

Two constants are important for the calculation of linkage data in *coli* crosses. The first constant (α) is an integration constant and its value is equal to the probability that a zygote will give rise to a recombinant of the desired selected type. It is also equal to the linkage frequency of a marker at genetic equilibrium. The incorporation frequencies are significantly different from cross to cross. This phenomenon will be discussed in our next paper³. The second constant (k) gives the average number of breakage events per unit of length.

Several additional features of the recombination process can be deduced from the experiments presented in this paper.

(1) This distance between origin and unselected marker has no influence upon the linkage frequency of the unselected marker as may be seen from the absence of this distance in Eqn. 2. This fact explains why markers which are located near the origin are incorporated with the same efficiency as more distal markers¹⁴. Consequently, the point of the acceptor chromosome corresponding with the origin of the Hfr is not a point of counter-selection as was previously suggested¹⁷ (compare also ref. 3).

(2) The distance between the selected marker and the distal end of the Hfr chromosome is also absent from Eqn. 2. This explains the absence of any influence of the length of the distal segment on the incorporation of a selected marker in a

blender experiment². The absence of the lengths of proximal and distal segments in Eqn. 2 suggests that the assumption of obligatory breakage events at both ends of the Hfr fragment is correct.

(3) The average number of breakage events (k) per min transfer time is constant over the whole segment in a particular cross. The fact that a proximal and a more distal segment have the same breakage probability per unit length shows that the recombination process begins at a certain time after the completion of transfer. If the process started immediately after the beginning of transfer, a higher value of k for proximal segments is expected. The explanation is probably that effective pairing between donor fragment and acceptor chromosome begins after the completion of transfer, and that pairing is necessary for the initiation of the recombination process. The constant probability can also be explained by the alternative model given above.

(4) The breakage events are randomly distributed over the donor segment, showing that a high number of points is available for the breakage process. The actual number of breakage events is rather low. This implies that the probability of the breakage event is low. An explanation may be that a low number of enzyme molecules is available.

Unselected markers can be mapped with the method described in this paper when two unselected markers with known transfer time are included in the cross. The constants α and k can be calculated from the linkage data of the two markers and the distances in time units of other markers may then be determined by a plot as given in Fig. 4. Exact determination of α is unfortunately impossible. On the other hand, the method is based on transfer times of known markers and errors in the determination of transfer times makes an accurate determination of map distances impossible. The values of α and $k\mu$ can be calculated without making use of the transfer times of unselected markers. This method will be presented in the next paper³.

REFERENCES

- 1 ARBER, W. AND C. LATASDE-DOROLLE, Erweiterung des Wirtsbereiches des Bakteriophagen λ auf *Escherichia coli* B. *Pathol. Microbiol.*, 24 (1961) 1012-1018.
- 2 DE HAAN, P. G. AND J. D. GROSS, Transfer delay and chromosome withdrawal during conjugation in *E. coli*. *Genet. Res.*, 3 (1962) 251-272.
- 3 DE HAAN, P. G. AND C. VERHOEF, Genetic recombination in *E. coli*. II. Calculation of incorporation frequency and relative map distance by recombinant analysis. *Mutation Res.*, 3 (1966) 111-117.
- 4 FOX, M. S. AND R. D. HOTCHKISS, Fate of transforming deoxyribonucleate following fixation by transformable bacteria. *Nature*, 187 (1960) 1002-1006.
- 5 HAYES, W., *The Genetics of Bacteria and Their Viruses*, Blackwell Scientific Publications, Oxford, 1964, p. 587-590.
- 6 IHLER, G. AND M. MESELSON, Genetic recombination in bacteriophage λ by breakage and joining of DNA molecules. *Virology*, 21 (1963) 7-10.
- 7 JACOB, F. AND E. L. WOLLMAN, Étude génétique d'un bactériophage tempéré d'*Escherichia coli*. III. Effet du rayonnement ultraviolet sur la recombinaison génétique. *Ann. Inst. Pasteur*, 88 (1955) 724-749.
- 8 JACOB, F. AND E. L. WOLLMAN, Sur le processus de conjugaison et de recombinaison chez *Escherichia coli*. I. L'induction par conjugaison ou induction zygotique. *Ann. Inst. Pasteur*, 91 (1956) 486-515.
- 9 JACOB, F. AND E. L. WOLLMAN, *Sexuality and the Genetics of Bacteria*. Academic Press, New York, 1961, p. 227.
- 10 KELLENBERGER, G., M. L. ZICHICHI AND J. J. WEIGLE, Exchange of DNA in the recombination of bacteriophage λ . *Proc. Natl. Acad. Sci. U.S.*, 47 (1961) 869-878.

- 11 LEDERBERG, J., Recombination mechanisms in bacteria. *J. Cellular Comp. Physiol.*, 45, Suppl. 2 (1955) 75-108.
- 12 MESELSON, M. AND J. J. WEIGLE, Chromosome breakage accompanying genetic recombination in bacteriophage. *Proc. Natl. Acad. Sci. U.S.*, 47 (1961) 857-868.
- 13 MESELSON, M., On the mechanism of genetic recombination between DNA molecules. *J. Mol. Biol.*, 9 (1964) 734-745.
- 14 PITTARD, J. AND E. A. ADELBERG, Gene transfer by F' strains of *Escherichia coli* K₁₂. III. An analysis of the recombination events occurring in the F' male and in the zygotes. *Genetics*, 49 (1964) 995-1007.
- 15 REEVES, H., Role of Hfr mutants in F⁺ × F⁻ crosses in *E. coli* K₁₂. *Nature*, 185 (1960) 265-266.
- 16 SIDDIQI, O. H., Incorporation of parental DNA into genetic recombinants of *E. coli*. *Proc. Natl. Acad. Sci. U.S.*, 49 (1963) 589-592.
- 17 STOUTHAMER, A. H., P. G. DE HAAN AND H. J. J. NIJKAMP, Mapping of purine markers in *Escherichia coli* K₁₂. *Genet. Res.*, in the press.
- 18 TAYLOR, A. L. AND M. S. THOMAN, The genetic map of *E. coli* K₁₂. *Genetics*, 50 (1964) 659-677.
- 19 VOLL, M. J. AND S. H. GOODGAL, Recombination during transformation in *Hemophilus influenzae*. *Proc. Natl. Acad. Sci. U.S.*, 47 (1961) 505-512.

Mutation Res., 3 (1966) 101-110