A MODEL FOR THREE-POINT ANALYSIS OF RANDOM GENERAL TRANSDUCTION

T. T. WU

Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts

Received March 2, 1966

SINCE the discovery of bacterial transduction (LEDERBERG, LEDERBERG, ZINDER and LIVELY 1951), it has been recognized that there are two distinct classes: special (MORSE, LEDERBERG and LEDERBERG 1956) and general (ZINDER and LEDERBERG 1952; LENNOX 1955; and JACOB 1955), which involve different mechanisms. In the case of general transduction, the chromosome segments carried by the phages can be obtained either from breaking the bacterial chromosome at fixed points as in the phage P22-Salmonella system (ZINDER and LEDERBERG 1952), or from breaking it randomly as in the phage P1-*Escherichia coli* system (LENNOX 1955). In the present study, a model is proposed for the random general transduction.

The model: The three-point analysis of transduction has been used extensively for defining the linear order of genetic markers (YANOFSKY and LENNOX 1959; GROSS and ENGLESBERG 1959; ADLER and KAISER 1963; FRANKLIN and HOWARD 1965; GLANSDORFF 1965; SIGNER, BECKWITH and BRENNER 1965; and COZZARELLI and LIN 1966). Twelve measurements in three sets of fours are analyzed (Figure 1). One of the markers carried by the transducing segment is selected, and the other two markers are then analyzed for their origin from either the transduc-

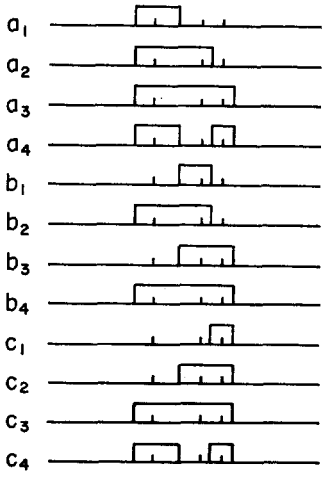

FIGURE 1.—The 12 measurements in three sets of four used in three-point analysis of transduction.

ing segment or the bacterial chromosome. Measurements a_4 and c_4 require four crossover events, so that they appear as small numbers. The order of the three markers can thus be established.

The remaining ten measurements are conveniently expressed as percentages, i.e.

$$L1 = a_1/(a_1 + a_2 + a_3), \tag{1}$$
$$L2 = a_2/(a_1 + a_2 + a_3), \tag{2}$$
$$L3 = a_3/(a_1 + a_2 + a_3), \tag{3}$$
$$M1 = b_1/(b_1 + b_2 + b_3 + b_4), \tag{4}$$
$$M2 = b_2/(b_1 + b_2 + b_3 + b_4), \tag{5}$$
$$M3 = b_3/(b_1 + b_2 + b_3 + b_4), \tag{6}$$
$$M4 = b_4/(b_1 + b_2 + b_3 + b_4), \tag{7}$$
$$N2 = c_1/(c_1 + c_2 + c_3), \tag{8}$$
$$N2 = c_2/(c_1 + c_2 + c_3), \tag{9}$$
and
$$N3 = c_3/(c_1 + c_2 + c_3). \tag{10}$$

The number a_1 can be expressed as (Figure 2), to the lowest order of μ,

$$a_1 = \mu^2 p \left[\int_0^{k-s} xs\, dx + \int_{k-s}^{k} x(k-x)dx \right], \tag{11}$$

where μ is the proportionality constant between the crossover frequency and the length of the segment of the chromosome involved, p is a distribution function of where the transducing particle is obtained from the bacterial chromosome, and can be taken as a constant with respect to the small segment involved in transduction for nearly random general transductions, k is the total length of the transducing particle and is assumed to be constant, s is the distance between the first and second markers (t is the distance between the second and third markers), and x is the length of the segment of the transducing particle to the left of the first marker. Thus, we have

$$a_1 = \mu^2 p\, [k^3 - (k-s)^3]/6. \tag{12}$$

A similar calculation gives that

$$a_2 = b_2 = \mu^2 p\, [(k-s)^3 - (k-s-t)^3]/6, \tag{13}$$
$$a_3 = b_4 = c_3 = \mu^2 p\, (k-s-t)^3/6, \tag{14}$$
$$b_1 = \mu^2 p\, st\, (2k-s-t)/2, \tag{15}$$
$$b_3 = c_2 = \mu^2 p\, [(k-t)^3 - (k-s-t)^3]/6, \tag{16}$$
and
$$c_1 = \mu^2 p\, [k^3 - (k-t)^3]/6. \tag{17}$$

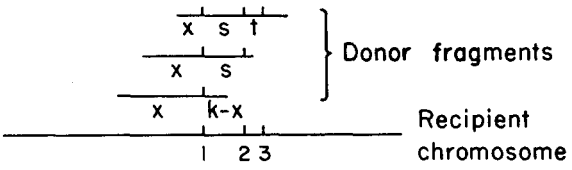

FIGURE 2.—The various possible locations of the donor fragment with respect to the recipient chromosome, so that the a_1 type of measurement can result.

To this order of μ, a_4 and c_4 can be neglected, and the effect of interference can be ignored. It can also be shown that

$$\begin{aligned} a_1 + a_2 + a_3 &= b_1 + b_2 + b_3 + b_4 \\ &= c_1 + c_2 + c_3 \\ &= \mu^2 pk^3/6. \end{aligned} \tag{18}$$

In other words, with $\alpha = s/k$ and $\beta = t/k$, equations (1) to (18) can be reduced to

$$A = L1 = 1 - (1 - \alpha)^3, \tag{19}$$

$$B = L2 = M2 = (1 - \alpha)^3 - (1 - \alpha - \beta)^3, \tag{20}$$

$$C = L3 = M4 = N3 = (1 - \alpha - \beta)^3, \tag{21}$$

$$D = M1 = 3\alpha\beta(2 - \alpha - \beta), \tag{22}$$

$$E = M3 = N2 = (1 - \beta)^3 - (1 - \alpha - \beta)^3, \tag{23}$$

and

$$F = N1 = 1 - (1 - \beta)^3, \tag{24}$$

where only α and β are independent.

TABLE 1

Predictions for su_c-tryC-pyrF, *based on equations (19)–(24)*

	A	*B*	*C*	*D*	*E*	*F*	α	β
Experimental	0.51	0.30	0.19	...	...	...	...	...
	...	0.20	0.17	0.20	0.43	...	...	...
	...	...	0.14	...	0.43	0.43	...	...
Calculated	...	0.25	0.24	0.17	0.34	...	0.21	0.17

TABLE 2

Predictions for su_c-tryA-pyrF, *based on equations (19)–(24)*

	A	*B*	*C*	*D*	*E*	*F*	α	β
Experimental	0.75	0.19	0.06	...	...	...	...	...
	...	0.18	0.10	0.31	0.41	...	...	...
	...	...	0.07	...	0.52	0.40	...	...
Calculated	...	0.14	0.11	0.26	0.49	...	0.37	0.16

TABLE 3

Predictions for purB-su_c-tryC, *based on equations (19)–(24)*

	A	*B*	*C*	*D*	*E*	*F*	α	β
Experimental	0.99	0.01	0.00	...	...	...	...	...
	...	0.01	0.00	0.44	0.55	...	...	...
	...	...	0.00	...	0.36	0.64	...	...
Calculated	...	0.01	0.00	0.62	0.37	...	0.74	0.26

TABLE 4

Predictions for arg-1-met-1-glpK, *based on equations (19)–(24)*

	A	*B*	*C*	*D*	*E*	*F*	α	β
Experimental	0.64	0.15	0.21	...	...	...	...	...
	...	0.22	0.19	0.19	0.41	...	...	...
	...	...	0.19	...	0.48	0.33	...	...
Calculated	...	0.16	0.20	0.17	0.47	...	0.29	0.13

Comparison with experiment: In order to test the the predictions of the present model and to exclude the applicability of the model proposed by DAWSON (1963) for the phage P22-Salmonella system, experimental results on markers with low cotransduction frequencies (SIGNER, BECKWITH and BRENNER 1965; and COZZARELLI and LIN 1966) are examined. In Tables 1 to 4, the first three lines give the experimental data, and the fourth lines show the calculated values of *B*, *C*, *D* and *E* from the given values of *A* and *F* using equations (19) to (24), together with the values of α and β. The theoretical predictions are in good agreement with the observed values. A comparison of the results given in Table 1 and 2 indicates the sequence of $su_c - tryC - eryA - pyrF$ (SIGNER, BECKWITH and BRENNER 1965), and the average of $\alpha + \beta$ in these two cases gives that the three intervals of the above sequence are roughly 25%, 6%, and 14% of the length of the transduced piece. The interval between su_c and *tryC* is further confirmed by the result given in Table 3.

DISCUSSION

The primary purpose of a model is to simplify experimental data and to make new predictions. The present one effectively reduces ten numbers to two. In addition, these two numbers, α and β, have direct biological significances, by giving estimates to distances between markers with respect to the length of the tranducing particles. The molecular weight of the tranducing particles of phage P1 has been estimated to be 6×10^7 daltons (IKEDA and TOMIZAWA 1965a), or approximately a few hundred cistrons. If the values of α and β are accurate to within a few percent, they can locate markers on the chromosome to within a few cistrons.

This prediction may then be used to bridge the gap between conjugation experiments and cotransduction studies. The interrupted conjugation technique (WOLLMAN, JACOB and HAYES 1956) has been used extensively to delineate the genetic map of *E. coli* (TAYLOR and THOMAN 1964). However, the locations of many markers are still uncertain. The relative positions of closely linked markers are precisely located by their high frequencies of cotransduction (YANOFSKY and LENNOX 1959). For not so closely linked markers, the ordering can be determined by three-point analysis, but the cotransduction frequencies are usually not accurate enough to locate the markers definitely. The present model provides a more precise scheme to assign the locations of markers that can be cotransduced. In fact, a similar analysis gives that

$$\text{cotransduction frequency} = \left(1 - \frac{\text{distance between two markers}}{\text{length of transducing particle}}\right)^3.$$

With accurate measurements of these frequencies (SIGNER, BECKWITH and BRENNER 1965), the distances between markers are indeed additive.

Finally, the two simplifying assumptions used in the model may be amended, if required. The breakage of the bacterial chromosome in the phage P1-*Escherichia coli* system may not be completely random, since in the phage P22-Salmonella system, not only is it broken at fixed points, but the frequencies of transduction also vary over different regions of the map (OZEKI 1959). The lengths of the transducing particles have been found to be heterogenous (IKEDA and TOMIZAWA 1965b). These two distribution functions can indeed be incorporated into the theory. Nevertheless, the available experimental data seem to indicate that such refinements may not be essential, unless more precise measurements become possible.

This work is supported by a grant to DR. E. C. C. LIN from the National Science Foundation (GB-3527).

SUMMARY

A model has been constructed so that three-point analysis of random general transduction can be used to calculate the distances between the markers with respect to the average length of the transducing particle. Predictions based on data from the phage P1-*Escherichia coli* system show good agreement with observed values.

LITERATURE CITED

ADLER, J., and A. D. KAISER, 1963 Mapping of the galactose genes of *Escherichia coli* by transduction with phage P1. Virology **19**: 117–126.

COZZARELLI, N. R., and E. C. C. LIN, 1966 The chromosomal location of the structure gene for glycerol kinase in *Escherichia coli*. J. Bact. **91**: 1763–1766.

DAWSON, G. W. P., 1963 The interpretation of data from transduction experiments. Genet. Res. **4**: 416–426.

FRANKLIN, N. C., and B. D. HOWARD, 1965 Double heterogenotes formed by P1-mediated transduction of *lac* genes in *Escherichia coli*. Virology **25**: 98–110.

GLANSDORFF, N., 1965 Topography of cotransducible mutations in *Escherichia coli* K-12. Genetics **51**: 167–179.

GROSS, J., and E. ENGLESBERG, 1959 Determination of the order of mutational sites governing L-arabinose utilization in *Escherichia coli* B/r by transduction with phage P1bt. Virology **9**: 314–331.

IKEDA, H., and J. TOMIZAWA, 1965a Transducing fragments in generalized transduction by phage P1. I. Molecular origin of the fragments. J. Mol. Biol. **14**: 85–109. —— 1965b Transducing fragments in generalized transduction by phage P1. III. Studies with small phage particles. J. Mol. Biol. **14**: 120–129.

JACOB, F., 1955 Transduction of lysogeny in *Escherichia coli*. Virology **1**: 207–220.

LEDERBERG, J., E. M. LEDERBERG, N. D. ZINDER, and E. R. LIVELY, 1951 Recombination analysis of bacterial heredity. Cold Spring Harbor Symp. Quant. Biol. **16**: 413–441.

LENNOX, E. S., 1955 Transduction of linked genetic characters of the host by bacteriophage P1. Virology **1**: 190–206.

MORSE, M. L., E. M. LEDERBERG, and J. LEDERBERG, 1956 Transduction in *Escherichia coli* K-12. Genetics **41**: 142–156.

OZEKI, H., 1959 Chromosome fragments participating in transduction in *Salmonella typhimurium*. Genetics **44**: 457–470.

SIGNER, E. R., J. R. BECKWITH, and S. BRENNER, 1965 Mapping of suppressor loci in *Escherichia coli*. J. Mol. Biol. **14**: 153–166.

TAYLOR, A. L., and M. S. THOMAN, 1964 The genetic map of *Escherichia coli* K-12. Genetics **50**: 659–677.

WOLLMAN, E. L., F. JACOB, and W. HAYES, 1956 Conjugation and genetic recombination in *Escherichia coli* K-12. Cold Spring Harbor Symp. Quant. Biol. **21**: 141–162.

YANOFSKY, C., and E. S. LENNOX, 1959 Transduction and recombination study of linkage relationships among the genes controlling tryptophan synthesis in *Escherichia coli*. Virology **8**: 425–447.

ZINDER, N. D., and J. LEDERBERG, 1952 Genetic exchange in Salmonella. J. Bacteriol. **64**: 679–699.