SPECIFICITY OF INSERTION BY THE TRANSLOCATABLE TETRACYCLINE-RESISTANCE ELEMENT Tn*10*

NANCY KLECKNER*†, DAVID A. STEELE*, KATHERINE REICHARDT† AND DAVID BOTSTEIN†

**Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138*

AND

†Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Manuscript received October 30, 1978
Revised Copy received February 5, 1979

ABSTRACT

Genetic analysis of 131 independent transpositions of the tetracycline-resistance element Tn*10* from a single site in phage P22 into the histidine operon of *Salmonella typhimurium* reveals that Tn*10* insertions are not randomly distributed along this chromosomal target. The insertions occur in 22 different "clusters"; insertions within each cluster are very tightly linked in recombination tests. Tn*10* insertions are not evenly distributed among the identified clusters. The existence of these clusters suggests that this chromosomal target contains particular genetic signals that guide Tn*10* to particular preferred positions for insertion. Insertions within each cluster occur in both orientations with roughly equal frequency.——The relationship among different insertions within each cluster has been examined. The resolution of genetic mapping places an upper limit of about 50 basepairs on the distance between different insertions within a cluster. Different insertions within a cluster usually have the same reversion frequency; however, heterogeneity in reversion frequency has been detected in at least two clusters. For most clusters, the available data are consistent with the simple possibility that all insertions within a cluster are at identical positions; however, the data do not exclude other possibilities.

PROKARYOTIC organisms contain a wide variety of translocatable genetic elements. The mechanisms by which these elements move from one place to another are not understood. Since each element translocates as a discrete, non-permuted linear segment of DNA, it seems clear that translocation involves very special sites at the ends of the element. Direct evidence that such sites are required has been obtained for one element, Tn*1* (HEFFRON *et al.* 1977). By contrast, the translocation process seems generally to be much less specific for particular sites on a DNA target. These and other properties of translocatable elements are reviewed in detail elsewhere (BUKHARI, SHAPIRO and ADHYA 1977; KLECKNER 1977; STARLINGER and SAEDLER 1976.)

The work described below concerns the mechanism of translocation of one

particular element, the tetracycline-resistance element Tn*10* (KLECKNER *et al.* 1975; BOTSTEIN and KLECKNER 1977). Tn*10* is 9300 basepairs in length, with 1400-basepair inverted repetitions at its ends. These repetitions may be intact IS-like sequences (CHAN 1974; SHARP, COHEN and DAVIDSON 1973; ROSS, SWAN and KLECKNER in preparation). The specificity of Tn*10* insertion for sites on a well-defined DNA target was examined by isolating and analyzing genetically 131 independent transpositions of Tn*10* from one site in bacteriophage P22 into the 10,000-basepair histidine operon of *Salmonella typhimurium*.

The results of such analysis, presented below, show that Tn*10* insertions are not randomly distributed along the histidine operon; rather, they occur as discrete clusters of very tightly linked insertions. Twenty-two such clusters have thus far been identified and mapped. Analogous results have been obtained by FOSTER (1977) for Tn*10* insertions in the *lacZ* gene. Tn*10* insertions are not evenly distributed among different clusters, but seem rather to occur preferentially in certain "hot spots". These and other results provide a detailed genetic picture of the interaction between Tn*10* and its target.

MATERIALS AND METHODS

Bacterial strains: All of the bacterial strains used in this work are derivatives of *Salmonella typhimurium* LT2. Tn*10* insertions were isolated in three strain backgrounds: wild-type LT2 (obtained from J. ROTH); NK80, a *his*+ *edd*⁻ derivative of wild-type LT2 constructed by transduction to *his*+ *gnd*+ of strain TR2932 (*hisO1242 his2236 gnd*- *edd*⁻) of J. ROTH; and DB47, a *recA1* prototrophic derivative of wild-type LT2 whose construction is described by BOTSTEIN and MATZ (1970). These strains are all closely related, although not perfectly isogenic.

F'ts114 *lac*+ episomes containing Tn*10* insertions in F (denoted *zzf*::Tn*10*) were used for chromosome mobilization experiments in Table 3. These episomes originated in a str^R *pyrC7* host. Strains carrying *zzf20*, *zzf21* and *zzf22* are TT627, TT628 and TT629, respectively, were provided by J. ROTH, and are described by KLECKNER, ROTH and BOTSTEIN 1977 and CHUMLEY, MENZEL and ROTH 1979. Episomes were introduced into *his*::Tn*10* strains by selecting for the Lac+ phenotype of the episome and counterselecting the donor by requiring Pyr+.

HisG-deletion strains used for mapping experiments in Table 2 and Figure 3 were obtained from J. ROTH and are described by HOPPE *et al.* 1978.

We are grateful to J. ROTH for providing us with the above strains and about half of the insertions analyzed in this study, which were isolated in the course of other experiments in his laboratory.

Bacteriophage strains: The P22 phage used for generalized transduction experiments was P22 *int3* HT12/4. The *int3* mutation (SMITH and LEVINE 1967) guarantees the recovery of nonlysogenic transductants by preventing phage integration; the HT12/4 mutation (SCHMEIGER 1972) increases the frequency of generalized transduction. The construction of this phage is described in KLECKNER *et al.* 1975.

The P22::Tn*10* phage used as a Tn*10* donor for isolation of all *his*::Tn*10* insertions used in this work is P22 Tc*10*::Tn*10 12*⁻ *amN11 13*⁻ *amH101 c2ts29 int3*. The construction and rationale for use of this vehicle is described by KLECKNER, ROTH and BOTSTEIN (1977). Lysates of this phage were made by induction of a suitable lysogen (NK337). This lysogen carries an amber suppressor (*su*+*19*) and an amber mutation in the leucine operon (*leuA414*). Lysates were made by growing NK337 to exponential phase in LB broth at 30°, raising the temperature to 40° and then incubating with aeration for three hr. The culture was then lysed with chloroform; debris was removed by centrifugation. Purified P22 baseplate parts were added as described in ISRAEL (1967). Phage particles collected by high-speed centrifugation, purified by centrifugation through a CsCl step gradient as described in BOTSTEIN (1968) and dialyzed to remove CsCl.

These lysates contain few particles capable of producing a plaque on single infection, since P22::Tn*10* genomes are oversize and because *int*⁻ lysogens produce few viable phage in any case (CHAN *et al.* 1972). Lysates do, however, contain normal numbers of particles, many of which carry the *Tn10*-containing portion of the P22-genome. Particle titer for these lysates was determined by measuring UV absorbance at 260 nm; $A_{260} = 1$ is equivalent to 5×10^{11} infective particles.

Media: LB medium (broth and agar) and M9 minimal agar are described by MILLER (1972). Green indicator agar is described by CHAN *et al.* (1972). Supplements were used routinely at the following concentrations: glucose, 0.2%; amino acids, 20 μg per ml; histidinol, 400 μg per ml; tetracycline hydrochloride (Sigma), 25 μg per ml; streptomycin, 500 μg per ml.

Green indicator plates containing tetracycline were used for the initial isolation of chromosomal Tn*10* insertions after infection of a bacterial host with the P22 phage vehicle. Two kinds of plates were used: green agar containing 25 μg per ml tetracycline and 0.01 M EGTA [Ethylenebis(oxyethylenenitrilo) tetraacetic acid; Eastman] or green agar plates containing 50 μg per ml tetracycline and no EGTA.

Isolation of independent his::*Tn*10 *insertions*: All insertions described in this paper were made by adsorbing a lysate of NK337 (P22 *12⁻ 13⁻ c2ts int⁻* Tn*10*; see Bacteriophage strains above) to cells of the desired bacterial host that had been grown to 5×10^8 per ml in LB broth. Mixtures were arranged so that the multiplicity of infection was approximately ten particles per infecting cell. After 30 min at 37°, each mixture was plated on green indicator plates containing tetracycline (see *Media* above). Plates were incubated at 40°; colonies were visible after 24 hr. After 36 hr, each plate was replica plated onto glucose minimal and LB agar plates, each containing tetracycline. Auxotrophs were identified by comparison of the two plates, and histidine auxotrophs were identified from these by further analysis. All of the His^-tet^R strains thus isolated contain Tn*10* insertions in the histidine operon by the criteria that His^+ revertants are tetracycline sensitive and that the His^- phenotype is 100% co-transducible with the tetracycline-resistance phenotype.

Since the conclusions from this analysis depend critically on the independent occurrence of each *his*::Tn*10* insertion studied, care was taken to insure that this requirement was met. Many independent mixtures of phage and cells were made, and only one *his*::Tn*10* insertion was saved from each adsorption mixture. If the translocation events giving rise to *his*::Tn*10* insertions occur after addition of phage to cells, this procedure guarantees that insertions from different adsorption mixtures arose independently. Since P22 is a generalized transducing phage, there was in principle the possibility that each of the insertions isolated actually derived from a translocation event that had occurred in the original NK337 lysogen and was simply being transferred to the new recipient by generalized transduction. If this were the case, insertions obtained from any one lysate would not be independent. In order to ensure that the nonrandom distribution of *his*::Tn*10* insertions obtained did not simply reflect the clonal origin of different insertions in the donor NK337 strain, we obtained insertions from separate lysates made from independent colony isolates of NK337, and we isolated approximately one third of all the insertions analyzed using a *recA*⁻ bacterial recipient. The insertions obtained under these various regimes were indistinguishable with respect to their distribution along the histidine operon; in all cases a large proportion of the insertions mapped at the highly preferred position at the beginning of *hisG*, and a comparison of the distribution of insertions obtained in *recA*⁺ and *recA*⁻ shows that insertions map at the same preferred position in both cases. We conclude that each of the insertions analyzed is the product of an independent translocation event.

The isolation of insertions from a number of independent NK337 lysates also guards against the possibility that any one lysate was for some reason atypical.

Transductional crosses: For all recombination tests, the following procedure was used. Recipient strains were grown to a density of 5×10^8 cells per ml, from single colonies, in LB broth. 10^8 recipient cells (0.2 ml) were mixed with 4×10^9 plaque-forming units of a P22 *int3* HT12/4 transducing lysate grown on the appropriate donor strain as described below. The mixture of phage and cells was usually made directly on the selective plates, and the mixture then spread. Plates were scored after incubation at 37° for 40 hr.

Growth of transducing lysates: A stock of P22 *int3* HT12/4 was grown in a strain carrying an extensive deletion of the histidine operon (NK455 = *his3050 edd*$^-$) according to the procedure of BOTSTEIN and MATZ (1970). This stock was used to prepare transducing lysates as follows: The desired bacterial strain was grown to 3×10^8 cells per ml in LB broth; P22 *int3* HT12/4 grown on NK455 was then added at a multiplicity of 0.8 phage per cell. The mixture was then aerated at 37° for 120 min, chloroform was added, and the lysate treated as a normal phage stock (BOTSTEIN and MATZ 1970).

Chromosome mobilization tests: A series of *his*::Tn*10*/F'ts114 *lac*$^+$ *zzf*::Tn*10* strains were constructed as described in *Bacterial strains* above. Each such strain was tested for the relative frequency with which it could mobilize either a *lys*$^+$ or a *trp*$^+$ marker by mating with a *lys*$^-$*trp*$^-$Sm^R recipient (NK686 = *lys553 trp125*::Tn*10 str*^{R} *spc*^{R}). Each *his*::Tn*10*/F::Tn*10* strain was grown to a single colony at 32° on minimal lactose histidine medium; such a single colony was then grown overnight at 42° in the same medium to select for F'ts *lac* integration. This overnight culture was then subcultured by diluting 1:100 in the same medium at 42° and grown to 2×10^8 per ml with very gentle shaking. A suitable aliquot was then mixed with a ten-fold excess of the *lys*$^-$*trp*$^-$ recipient, which had been grown to 5 to 8×10^8 per ml in LB broth with aeration. The mating mixture was incubated at 37° for three hr with gentle mixing and then diluted and plated on selective plates. Plates were incubated for several days at 37° and scored.

Calculation of reversion rates (Figure 4): Reversion frequencies per generation of *his*::Tn*10* insertions were calculated according to LURIA and DELBRUCK (1943) using a method that makes use of the average number of revertant bacteria per culture (r), the total number of bacteria per culture plated (N_t), and the number of independent cultures assayed (C). The relationship between (r) and (aN_t) is plotted for different values of (C) in Figure 1 of LURIA and DELBRUCK (1943). In the experiments described below, $C = 5$; in practice, the value of (r) is determined by finding the average number of revertant bacteria in a sample of a culture and multiplying this average per sample by the ratio of the total volume of culture to the volume of the sample by the ratio of the total value of culture to the volume of the sample. The values of reversion rates obtained by this method are in fact somewhat higher than the true values; however, this method is adequate to document the variation in reversion rates among different insertions.

RESULTS

Mapping of Tn10 insertions in the histidine operon: Tn*10* insertions in the histidine operon were identified among large numbers of auxotrophs generated by translocation of Tn*10* from the genome of bacteriophage P22 into the chromosome of *Salmonella typhimurium* LT2. The same P22::Tn*10* phage was used as donor for each of the insertions. Care was taken to insure that each insertion resulted from an independent translocation event; in addition, several different sets of insertions were isolated from several different lysates of this P22::Tn*10* phage. Insertions were isolated in both *recA*$^+$ and *recA*$^-$ bacterial hosts. The exact procedures for isolating independent *his*::Tn*10* insertions are described in the MATERIALS AND METHODS.

Each of the 131 *his*::Tn*10* insertions analyzed was first assigned to one of the known histidine structural genes by complementation tests against known *his*$^-$ mutants on F'*his* episomes. This procedure has been described previously (KLECKNER *et al.* 1975). The distribution of these insertions among the several genes in the operon is shown in Table 1. It reveals the nonrandomness with which Tn*10* chooses insertion positions. For example, 57 insertions were isolated in the *hisG* gene, a region of approximately 900 basepairs, while only five inser-

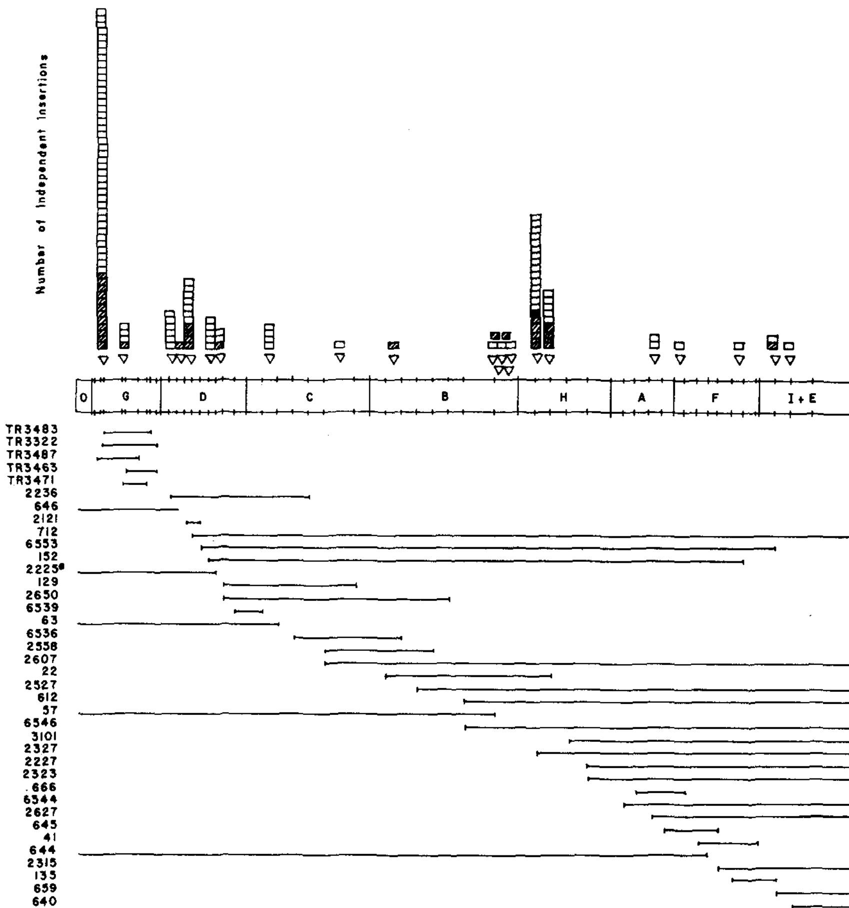

FIGURE 1.—Distribution of *his*::Tn*10* insertions among genetically defined clusters. Each ▽ represents a single cluster located in the indicated deletion interval. □ = insertion isolated in *recA*+ host; ▨ = insertion isolated in *recA*- host.

*Data for deletion 2225 are from J. ROTH (personal communication).

tions were isolated in the *hisB* gene, a region of about 1900 basepairs (BRENNER and AMES 1971).

Insertions within each gene were then mapped against one another by pairwise crosses using P22-mediated generalized transduction. For insertions in most genes, all possible combinations of pairwise crosses among different insertions were performed. This approach was impractical in genes containing many insertions. In these cases, each insertion was crossed against a representative subset of insertions in the same gene. Representative data are shown in Table 2A.

These pairwise crosses distributed all of the 131 insertions into 22 groups. Insertions within each group fail to recombine with each other to give *his*+ recombinants above the reversion level of the recipient strain; insertions in any one group will give *his*+ recombinants with all insertions in other groups. Each

TABLE 1

Distribution of Tn10 insertions among genes in the histidine operon

Gene	Size of gene* (basepairs)	Number of Tn10 insertions† recA+	recA−	Total
O/P	200	0	0	0
G	900	44	13	57
D	1100	20	6	26
C	1600	5	0	5
B	1900	3	3	6
H	1200	20	10	30
A	800	2	0	2
F	1100	2	0	2
I+E	1200	2	1	3
Total	10,000	98	32	131

* Based on estimates by BRENNER and AMES (1971).
† Insertions divided according to whether they were isolated in a *recA+* or *recA−* host.

TABLE 2

Recombination tests: hisG::*Tn*10 × hisG *deletion and* hisG::*Tn*10 × hisG::*Tn*10

Recipient:	Donor*:	Number of *his+* colonies *hisG9424* ::Tn*10*	*hisG9464* ::Tn*10*	*hisG9436* ::Tn*10*	*hisG9463* ::Tn*10*	*hisOGDCBHAFIE* deletion 3050†
A. *hisG*::Tn*10* insertions						
hisG9458		13	5	322	570	23
hisG9461		1	3	214	182	3
hisG9462		2	4	303	344	4
hisG9463		361	215	1	0	0
hisG9438		0	1	303	607	1
hisG9455		14	11	217	237	9
hisG9443		5	1	96	398	10
hisG9445		3	0	171	281	2
hisG8666		0	0	155	91	1
B. deletions‡						
hisG8522		51,44	30,27	0,0	1,1	0,0
hisG8476		0,0	0,0	0,0	0,0	0,0
hisG8526		0,0	0,0	0,0	0,0	0,0

* Transductional crosses were performed as follows: Recipients strains were grown to a density of 5×10^8 cells per ml from single colonies in LB broth. 10^8 recipient cells (0.2 ml) were mixed with 4×10^9 plaque-forming units of a transducing lysate of P22 *int3* HT12/4 grown on the appropriate donor strain, as described in MATERIALS AND METHODS. The mixture was usually made directly on the selective plate and then spread. Selective plates were M9 salts supplemented with glucose. Plates were incubated at 37° for 48 hr and then scored.
† *His*-deletion3050 is a deletion of the entire histidine operon and adjacent regions. Transductions using this deletion as a donor are included as controls for determining the level of revertant colonies in each recipient.
‡ Endpoints of these deletions are indicated in Figure 2. *hisG8522* is the closest deletion with which *hisG9424* and *hisG9462* (and other deletions in that cluster) still recombine; *hisG8476* is the next longest deletion, extending just past the insertions at this cluster. *hisG8526* is another deletion that extends through the positions of all tested insertions in *hisG*.

group was then mapped relative to the endpoints of known deletions in the histidine operon by analogous recombination tests. The distribution of insertions among these 22 groups and the locations of the groups along the operon are shown in Figure 1.

For convenience, each group will be referred to as a "cluster". The resolution of the pairwise crosses that define these clusters is described below. The most striking feature of the distribution shown in Figure 1, is that insertions are not evenly distributed among the identified clusters. It would appear that some positions might represent a kind of "hot spot" for Tn*10* insertion.

The reality of the proposed clusters was verified by fine-structure mapping of insertions in the *hisG* gene, taking advantage of the detailed genetic map of *hisG* constructed by HOPPE *et al.* (1979). This map is based on a collection of nearly 100 point mutations and 80 deletions; the map defines 40 deletion intervals in *hisG*. Crosses were performed between each of the *hisG*::Tn*10* insertions and appropriate *hisG* deletions; the results are shown in Figure 2. Representative data are given in Table 2B. Each of the insertions maps unambiguously into one of two nonadjacent deletion intervals. The resulting two clusters of insertions are the same two clusters previously defined by pairwise crosses among insertions.

Two other comments can be made about the distribution of Tn*10* insertions shown in Figure 1. First, insertions isolated in a *recA*⁻ host generally occur in the same clusters as insertions isolated in *recA*⁺ hosts. Second, the clusters themselves may not be randomly distributed along the histidine operon. Fourteen of 22 recombination groups lie in one of two regions, either in the *hisG/hisD* region or in a region including the end of gene *B* and the beginning of gene *H*.

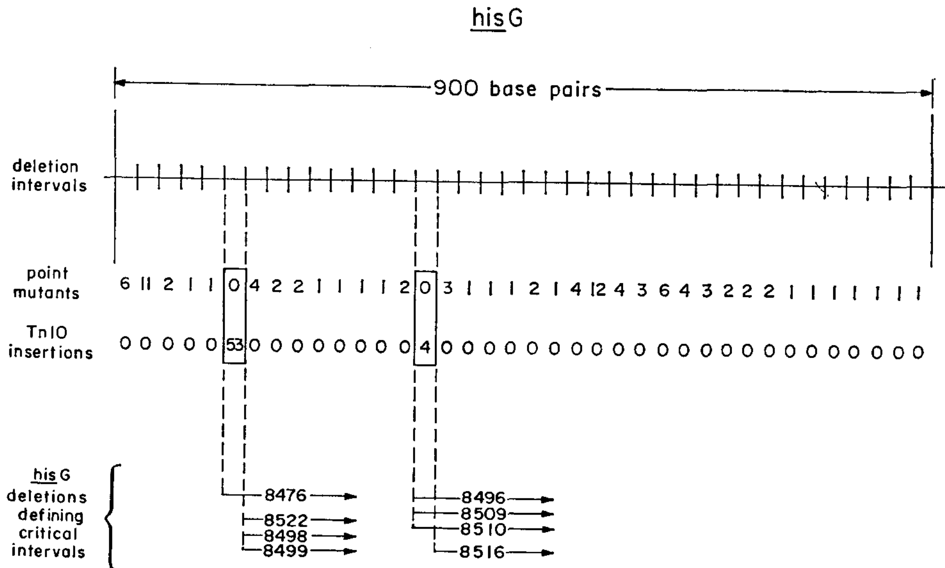

FIGURE 2.—Distribution of point mutations and Tn*10* insertions among the deletion intervals in *hisG* as defined by HOPPE *et al.* (1979).

*Orientations of Tn*10 *insertions at preferred positions*: The mapping data presented suggest that Tn*10* insertions in the histidine operon occur in clusters of very tightly linked insertions. In order to probe the relationships among independent insertions within each cluster, we determined the relative orientations of Tn*10* insertions in seven of the identified clusters by a genetic technique diagrammed in Figure 3 and described in further detail elsewhere (KLECKNER, ROTH and BOTSTEIN 1977; CHUMLEY, MENZEL and ROTH 1979.

When an F factor carries a sizeable stretch of sequences homologous to sequences on the bacterial chromosome, Hfr's will preferentially be formed by integration of F *via* general recombination between these two stretches of homology (BECKWITH, SIGNER and EPSTEIN 1966). Thus, when one Tn*10* element is present on F and a second element is present in the bacterial chromosome, Hfr's will preferentially arise by recombination between the homologous Tn*10* sequences. These Hfr's will have their origin of transfer at the position of the original Tn*10* insertion in the bacterial chromosome, and the direction of transfer will reflect the relative orientations of the two Tn*10* insertions involved. Integration of a given F′::Tn*10* derivative into each of two different *his*::Tn*10* strains will produce Hfr's having the same direction of transfer if the two *his*::Tn*10* insertions lie in the same orientation, and Hfr's having different directions of transfer if the Tn*10* insertions are in opposite orientations.

In practice, the direction of transfer of Hfr's resulting from F′ integration *via* Tn*10* homology is assessed by measuring the frequency with which markers on

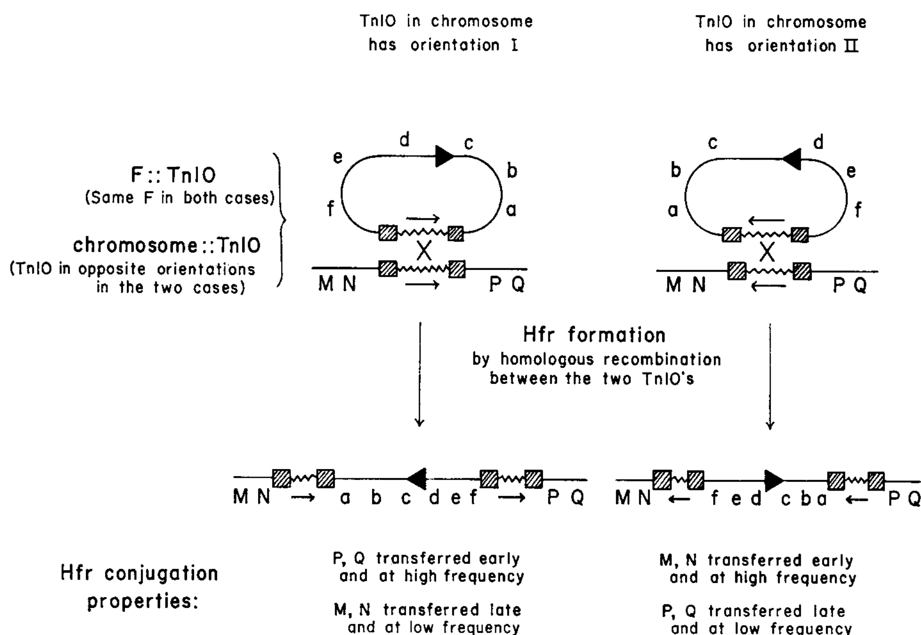

FIGURE 3.—Genetic determination of orientation of a chromosomal Tn*10* insertion by means of chromosome mobilization with F′ts114::Tn*10*.

either side of the chromosomal Tn*10* insertion are transferred by a cell population containing many independently arising integrated F′ derivatives. Such a population is obtained by using an Fts114 *lac*$^+$ episome carrying Tn*10*. Since Fts114 episomes are thermosensitive for F replication, one can select directly for derivatives in which the F′ episome is integrated into the bacterial chromosome simply by growing a culture of the desired *his*::Tn*10*/F′ts114 *lac*$^+$ *zzf*::Tn*10* strain at the nonpermissive temperature under conditions where the Lac$^+$ phenotype must be maintained (*i.e.*, growth on minimal lactose histidine medium). Such a preselected culture is then used as the donor population in a conjugational mating.

This method was used to determine the orientations of 46 independently isolated *his*::Tn*10* insertions in the *G*, *D*, and *C* genes. These insertions included 12 from the highly favored cluster in *hisG* and all of the insertions at other positions in these genes. Quantitative results for a number of *hisG*::Tn*10* insertions are shown in Table 3, and the results for each of the insertions tested are summarized in Table 4. The conclusion is clear: for every cluster containing more than one insertion, insertions were found in both orientations.

Reversions of his::*Tn*10 *insertions: genetic heterogeneity among insertions within certain clusters*: A further probe into the relationship among Tn*10* insertions within a cluster is provided by reversion analysis. Each of the 131 *his*::Tn*10* insertions analyzed thus far gives rise to *his*$^+$ revertants; reversion is accompanied by loss of the inserted Tn*10* element. This result further substantiates our previous conclusions (KLECKNER *et al.* 1975) that insertion of Tn*10* is usually quite precise with respect to the DNA target, since the interrupted structural gene can still be reconstructed in functional form in each of the 131 insertions.

Insertions in different clusters can give rise to revertants at frequencies per generation varying from about 10^{-6} to less than 10^{-10}; the distribution of reversion frequencies (calculated as described in MATERIALS AND METHODS) for 37

TABLE 3

Orientations of hisG::*Tn*10 *insertions*

hisG::Tn*10* insertion in chromosome / Tn*10* insertion on episome:	*zzf21*	*zzf22*	*zzf20*	Deduced relative orientations of *hisG*::Tn*10* insertions
	Relative frequencies of chromosome mobilization (*lys*$^+$ exconjugants/*trp*$^+$ exconjugants)			
hisG9445	0.046	19.0	0	B
hisG9459	0.041	20.0	0	B
hisG9427	0.098	19.9	11.3	B
hisG9455	0.14	16.2	11.7	B
hisG9462	0.12	12.5	5	B
hisG9461	0.03	13.2	$	B
hisG9425	0.75	9.9	5	B
hisG9438	3.8	0.025	0.11	A
hisG9458	5.0	0.058	0.15	A
hisG9468	6.9	0.03	0.11	A
hisG9424	3.7	0.05	0.13	A
hisG9443	3.0	0.04	0.085	A

TABLE 4

Orientations of Tn10 insertions from clusters in hisG, hisD *and* hisC

Cluster*	Number of insertions in Orientation A	Number of insertions in Orientation B	Total number of insertions tested
G1	5	7	12
G2	1	3	4
D1	3	1	4
D2	4	7	11
D3	3	3	6
D4	2	1	3
D5	1	0	1
C1	1	3	4
C2	0	1	1
Total	22	24	46

* Clusters as identified in Figure 1 are numbered in numerical order from left to right within each gene. All of the insertions at every cluster in *hisG*, *hisD* and *hisC* were tested for orientation by experiments identical to those shown in Table 3, except that only 12 insertions in the *G1* cluster were tested.

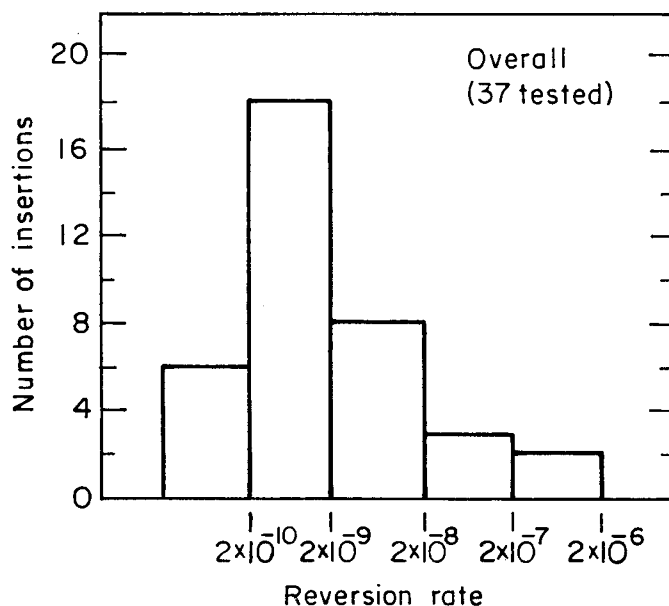

FIGURE 4.—Reversion rates of *his*::Tn*10* insertions. A reversion rate for each insertion was determined as follows: five independent clones of each insertion were grown from single colonies on LB agar to liquid cultures each containing about 6×10^9 total cells. Approximately 1.2×10^8 cells from each culture were spread on minimal glucose plates and incubated at 37°. Revertant colonies were counted after 40 hours. Reversion rates were calculated as described in MATERIALS AND METHODS using a method of LURIA and DELBRUCK (1943) that makes use of the average number of revertant bacteria per culture. The values of reversion rates obtained by this method are in fact somewhat higher than the true values; however, this method is adequate to assess differences in reversion rates among different strains.

representative insertions is shown in Figure 4. There is no simple correlation between the frequency of insertion into and the frequency of excision out of a given position. For example, most insertions that occupy unique positions revert at very low rates (2×10^{-9} or less), but several such insertions revert at extraordinarily high rates (10^{-5} to 10^{-6}). On the other hand, if one eliminates these "super-reverters" from consideration, insertions at the more frequently represented positions tend to revert somewhat more readily (in the range of 10^{-7} to 10^{-8}) than do insertions at positions represented by only one or two insertions (2×10^{-9} or less).

There is generally very little variation among reversion frequencies for insertions within a single cluster. We have compared the reversion properties of insertions within each of the clusters in *hisG*, *hisD*, and *hisC*. In this case, 15 independent clones of each insertion were scored for the number of revertants in 0.2 ml of a saturated broth culture. For five of the seven clusters tested, the different insertions within each cluster were indistinguishable in their reversion properties; in most cases differences in revertant frequencies of three-fold would have been detected. For two of the clusters, however, significant heterogeneity of revertant frequencies was seen. Data for one such cluster, the highly favored cluster in *hisG*, is shown in Figure 5. Although it is not possible to decide if the insertions tested fall into two categories or more, it is clear that different insertions

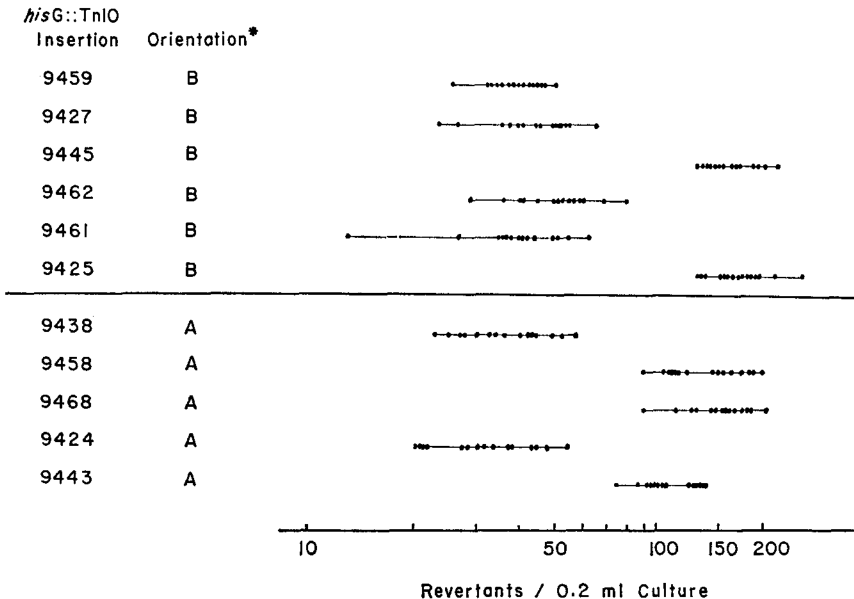

FIGURE 5.—Genetic heterogeneity among *hisG*::Tn*10* insertions in one cluster. Number of revertant colonies obtained from each of 15 independent clones of each *hisG*::Tn*10* insertion are plotted. Each number represents the number of revertant colonies obtained by spreading 0.2 ml of a saturated culture on a minimal glucose plate and incubating plates for five days at 37°. All of the cultures whose revertants are shown in this Figure were grown and plated in parallel in one experiment.

differ in their reversion properties; frequencies of revertants differ by as much as five-fold.

Experiments equivalent to that shown in Figure 5 have been carried out in several different strain backgrounds, including both *recA*$^+$ and *recA*$^-$ hosts. These experiments (data not shown) demonstrate that the characteristic differences in revertant frequency among insertions within a cluster are maintained in the several different strain backgrounds, although the absolute values of these frequencies may differ slightly from one background to another. Since insertions were moved from one strain to another by P22-mediated transduction, these results suggest that the heterogeneities in revertant frequencies observed among different insertions within a cluster are a property of the *his*::Tn*10* insertions themselves, rather than of some extraneous unlinked variation among the different strains tested.

DISCUSSION

One hundred thirty-one independent transpositions of Tn*10* from one site in bacteriophage P22 into the histidine operon of *Salmonella typhimurium* were subjected to genetic analysis. Tn*10* insertions are distributed nonrandomly along this target; insertions are clustered into groups of tightly-linked insertions. Twenty-two such "clusters" of insertions could be distinguished by recombination tests. Furthermore, Tn*10* insertions are not evenly distributed among the identified clusters; for example, one cluster contains roughly a third of all insertions, whereas there are 11 "clusters" represented by only one insertion. Comparisons among different insertions showed that, within each cluster, insertions occur in both orientations with roughly equal frequencies. In a few cases, genetic heterogeneity could be detected among insertions in a single orientation within a cluster, as assessed by differences in revertant frequencies.

Genetic size of clusters: The resolution of the genetic mapping methods used places an upper limit on the size of the DNA region defined by a single cluster. The *hisG* gene is approximately 900 basepairs long (BRENNER and AMES 1971), and the average length of a deletion interval is thus 900/40, or about 22 basepairs. The actual length of a given deletion interval may, of course, be quite different from this average length, either because of the statistical variation in actual lengths about the average, because deletion endpoints are clustered at a few particular positions along the physical map (as shown for deletions in the *gal* operon and *lacI* gene of *E. coli*; PFEIFER, HABERMANN and KUBAI-MARONI 1974; SCHMEISSNER, GANEN and MILLER 1977), or because recombination frequencies are uneven along the gene. The relatively even distribution of point mutations among the 40 deletion intervals in *hisG* (Figure 2 and HOPPE *et al.*, 1979) argues against the idea that any of these factors has produced major distortions in the relationship between genetic and physical distances. Most strikingly, both clusters of Tn*10* insertions actually occur in deletion intervals in which no point mutations occur, and within regions of the map where there are generally fewer point mutations per deletion interval, rather than more. If anything, these facts would argue that Tn*10* insertions lie in deletion intervals that are smaller than the

average, that is, that there are if anything more deletion endpoints per unit of physical distance in the regions of the Tn*10* insertions than elsewhere in the gene. These considerations suggest that, at the very least, the sizes of the deletion intervals containing Tn*10* clusters are unlikely to be grossly larger than the average.

Recombination between Tn*10* insertions and the closest deletion endpoint is readily detected. This suggests that the mapping method can conservatively be estimated to resolve (that is, detect recombinants between) insertions and deletions whose endpoints are as few as 50 basepairs apart. If crosses between two Tn*10* insertions are comparable to crosses between an insertion and a deletion, this estimate can be extended to the series of pairwise crosses carried out among differsent insertions in each gene. Thus, the data suggest that insertions within each cluster, which by definition are insertions that fail to recombine with one another, are probably not separated from one another by more than about 50 basepairs. These data do not, of course, exclude the most extreme possibility that the different insertions within each cluster are at exactly identical positions at the nucleotide level.

Significance of clusters: Clusters are groups of insertions that lie so close together that they are not separated by recombination tests. The occurrence of Tn*10* insertions in such clusters implies that Tn*10* inserts preferentially at certain highly preferred positions in the histidine operon, and in this sense the clusters define "hot spots" or "hot regions" for Tn*10* insertion. The occurrence of clusters does not preclude the possibility that Tn*10* can, in fact, insert at virtually any position in DNA with some measurable frequency; 11 of 131 insertions map alone at 11 distinct positions.

The occurrence of Tn*10* insertions in clusters must also mean that the histidine operon contains a series of genetically important sites or signals that influence Tn*10* insertion and play some role in guiding Tn*10* to particular positions for insertion. These sites or signals could in principle be involved only in the recognition between Tn*10* and its target and may, or may not, also be the actual physical site of ensuing integration events.

Since Tn*10* insertions are not evenly distributed among the identified clusters, perhaps some "guiding signals" are recognized or interacted with more efficiently than others.

Orientations of insertions: Insertions within each cluster occur in both possible orientations. This result suggests that Tn*10* can insert in either orientation in response to a given "guiding signal" on the target DNA. This implication will pose serious constraints on detailed molecular explanations of how Tn*10* specificity arises during interaction of Tn*10* with its target. The finding that insertion is fundamentally symmetrical with respect to the Tn*10* element itself is not unreasonable, in view of the fact that Tn*10* has a 1400-basepair inverted repetition at its ends.

The observation that Tn*10* insertions appear to occur in both possible orientations within a cluster is subject to one possible objection. Since Tn*10* has an inverted repetition at its ends, homologous recombination between these repeats

will result in the inversion of the intervening nonrepeated Tn*10* segment. Since it is the orientation of this segment that is assayed in our chromosome mobilization test, frequent recombination events of this type could randomize the apparent orientations of Tn*10* insertions and thus obscure any underlying asymmetry.

Several lines of evidence argue against this possibility. Inversion of Tn*10* during preparation of a donor P22::Tn*10* lysate should be detectable by cleavage of P22::Tn*10* DNA with a restiction endonuclease that does not cleave within the inverted repeat; such derivatives should give rise to "extra" Tn*10*-containing fragments. No such anomalous cleavage fragments have been seen in numerous digests of P22::Tn*10*; had more than 5% of the genomes suffered an inversion event, the resulting fragments would have been detected (G. WEINSTOCK, personal communication). The failure of chromosomal Tn*10* insertions to be randomized is best shown by the simple fact that one can determine unambiguously an orientation for each insertion tested. Furthermore, isolation of hundreds of independently occurring stable Hfr derivatives derived by integration of F::Tn*10* into a chromosomal Tn*10* insertion has failed to turn up a single Tn*10* inversion (R. MENZEL, F. CHUMLEY, M. JOHNSTON and J. ROTH, personal communication). The absence of extensive apparent inversion of Tn*10* is further substantiated by the fact that many of the phage::Tn*10* derivatives and many of the chromosomal *his*::Tn*10* insertions have been sub-cloned repeatedly and subsequently analyzed; no example of an inversion between the inverted repeats has been found. If recombination-promoted inversion between the ends of Tn*10* does occur, it does so at a rate too low to interfere with interpretation of our orientation data.

Heterogeneity in reversion properties: The observation that different Tn*10* insertions at different places in the histidine operon exhibit different reversion rates suggests that variation in reversion frequency reflects variation in the local environments of different insertions. Conversely, the observation that different insertions within a single cluster give revertants at different frequencies suggests that the different insertions are in different local environments; that is, they are not identical. The heterogeneity observed among different insertions within two clusters in the histidine operon is particularly significant because it is more often the case that insertions within a cluster give revertants at the same frequency, while characteristic frequencies vary from cluster to cluster.

The occurrence of Tn*10* insertions in clusters implies the presence along the target histidine operon of genetically important sites or signals that guide Tn*10* to particular preferred positions for insertion. In this context, the observation of genetic heterogeneity among tightly linked insertions is formally subject to two types of interpretation: either the heterogeneous cluster is actually a position in DNA where two (or more) important sites occur too close together to have been resolved by recombination tests, or else insertions that occur under the influence of a single site or signal are in fact not all identical. It is of course also possible that both situations obtain in certain instances; that is, two important sites are too tightly linked to be resolved and insertions that occur under the influence of a single site are not identical. Further experiments are required to determine

which of these possible interpretations accounts for the observed reversion heterogeneity.

Insertion specificities of other translocatable elements: A comparison of the insertion specificities of a number of translocatable elements is given with appropriate references in Table 5. Such comparisons are potentially difficult because insertion of any one element is generally examined using one or a few well-defined targets. Since different targets may contain different arrays of different "insertion sites" or other genetically important signals, the pattern observed may depend as much on the particular target chosen for study as on the element in question. Also, for the small IS sequences, several copies of which are present in the bacterial genome, the set of insertions isolated in a new target reflects the composite behavior of several different individual elements coming from several different donor positions.

Despite these potential complexities, one comes away with the general impression that the specificities of insertion of IS*1* (STARLINGER and SAEDLER 1976), IS*4* (PFEIFER, HABERMANN and KUBAI-MARONI 1977), Tn*10* (this work and FOSTER 1977), and of bacteriophage lambda when it inserts into secondary sites (SHIMADA, WEISBERG and GOTTESMAN 1973) are roughly comparable. That is,

TABLE 5

Insertion specificities of translocatable elements

Element	Target	Approx. target size (bp)	No. insertions examined	No. different sites	Specificity?	References
Tn*10*	Salmonella *his*	10,000	131	22	Yes	This work
Lambda*	*E. coli gal*		8	1	Yes	SHIMADA *et al.* 1973
	E. coli leu		2	1	Yes	SHIMADA *et al.* 1973
	E. coli ara	2100	4	3	Yes	BOULTER and LEE 1973
IS*1*	*E. coli gal* OP,T	2000	6	3	Yes	PFEIFER *et al.* 1977; STARLINGER and SAEDLER 1976
IS*4*†	*E. coli gal* OP,T	2000	41	1	Yes	PFEIFER *et al.* 1977
Mu-*1*	*E. coli lacZ*	3500	75	⩾50	Little/none	BUKHARI and ZIPSER 1972; DANIELL *et al.* 1972
Tn*5*	*E. coli lacZ*	3500	19	19	Little/none	BERG 1977
Tn*1*	P22 *ant*	800	21	⩾13	Regional§	WEINSTOCK 1977; WEINSTOCK and BOTSTEIN 1978
Tn*2*	RSF1010	<8300‡	38	⩾19	Little/none	HEFFRON *et al.* 1977; REUBENS *et al.* 1976
Tn*3*	R6–5	100,000	23	many	Regional§	KRETSCHMER and COHEN 1977
Tn*7*	RP4	52,000	36	⩾29	Regional§	BARTH and GRINTNER 1977

* These data are for lambda integration at positions other than the normal bacterial attachment site and a small number of very highly favored secondary positions.

† IS*4* has been defined only by the occurrence of a particular sequence inserted at a particular site in *gal*T.

‡ These targets are small plasmids; it is not known accurately what proportion of the sequences on these plasmids is essential for plasmid maintenance and replication.

§ Insertions occur preferentially in certain 1000–3000 basepair regions of the target; within these preferred regions, however, insertions occur at very many different positions.

when one examines insertions of any of these elements into a defined target of 3000 to 10,000 basepairs, one finds insertions at more than one position, but one also finds insertions that appear to map at the same position (generally indicated by their failure to recombine) even when the total number of insertions examined is relatively small (of the order of ten).

The relationships among different insertions within a genetically defined cluster has also been examined in some cases. For IS*1* insertions clustered in the *galOP* region, existing genetic and physical data suggest that the different insertions in the cluster differ in position by a few basepairs, leading to the suggestion (STARLINGER and SAEDLER 1976) that the integration mechanism of IS*1* might resemble that of type-I restriction enzymes (NATHANS and SMITH 1975) in which the recognition site is specified by the nucleotide sequence, but the cleavage site is not. Within this cluster in *galOP*, insertions occur in both orientations. Insertions of IS*4* have thus far been isolated only in a single deletion interval in *galT*. Insertions within this deletion interval also do not recombine; once again, however, insertions occur in both possible orientations at this position (PFEIFER, HABERMANN and KUBAI-MARONI 1977). These results for IS elements differ strikingly from the corresponding results for integration of lambda at secondary sites. In the case of lambda integration, insertions of lambda at any one secondary site occur in only one orientation (SHIMADA, WEISBERG and GOTTESMAN 1975; R. WEISBERG, personal communication). The significance of this difference remains to be determined.

The insertion patterns of the elements discussed above is quite different from those of several other elements, notably the related elements Tn*1*, *2*, *3* and *4*; Tn5, Tn7 and bacteriophage Mu-*1*. For the cases of Tn*1*,*2*,*3*,*4* and *7*, when insertions into a very large target (10,000 to 50,000 basepairs) are examined, insertions occur preferentially in particular regions of the target. These preferred regions are large, one to several thousand basepairs in length. Within these preferred regions, however, there is very little preference on the part of any of these elements for particular positions; there are virtually as many insertion positions as there are insertions (HEFFRON *et al.* 1977; WEINSTOCK 1977; WEINSTOCK and BOTSTEIN 1978; BARTH and GRINTNER 1977; KRETSCHMER and COHEN 1977) Bacteriophage Mu-*1* also appears to insert randomly, at least into the *lacZ* gene of *E. coli* (BUKHARI and ZIPSER 1972; DANIELL, ROBERTS and ABELSON 1972), and Tn5 inserts at many different places in the *lacZ* gene as well (BERG 1977).

Other factors influencing insertion: For the elements described above that exhibit regional specificity of insertion, there are clearly factors others than the nucleotide sequence in the immediate location of the integration position that influence where the element inserts. The seemingly nonrandom distribution of Tn*10* insertion clusters along the histidine operon carries a similar implication, although the factor(s) responsible might be very different in different cases.

We thank J. ROTH, I. HOPPE and M. JOHNSTON for numerous strains and generous access to unpublished results, including early indications of a preferred site of Tn*10* insertion in *hisG*. We thank ELEANOR STOCHAJ for excellent technical assistancee. N. KLECKNER was supported in part by an Public Health Service post-doctoral fellowship (AI05109), and in part by a

National Science Foundation Grant (83064). D. BOTSTEIN was the recipient of a Public Health Service Career Development Award (GM70325). The research was supported by Public Health Service grants (GM18973 and GM21253) and a National Science Foundation Grant (83064).

LITERATURE CITED

BARTH, P. T. and N. J. GRINTNER, 1977 Map of plasmid RP4 derived by insertion of transposon C. J. Mol. Biol. **113**: 455–474.

BECKWITH, J. R., E. R. SIGNER and W. EPSTEIN, 1966 Transposition of the *lac* region of *E. coli.* Cold Spring Harbor Symp. Quant. Biol. **31**: 393–401.

BERG, D., 1977 Insertion and excision of the transposable kanamycin resistance determinant Tn5. pp. 205–212. In: *DNA Insertion Elements, Plasmids and Episomes.* Edited by A. I. BUKHARI, J. A. SHAPIRO and S. ADHYA, Cold Spring Harbor Lab., Cold Spring Harbor, New York.

BOTSTEIN, D., 1968 Synthesis and maturation of phage P22 DNA. I. Identification of intermediates. J. Mol. Biol. **34**: 621–641.

BOTSTEIN, D. and N. KLECKNER, 1977 Translocation and illegitimate recombination by the tetracycline-resistance element Tn*10*. pp. 185–204. In: *DNA Insertion Elements, Plasmids, and Episomes.* Edited by A. I. BUKHARI, J. A. SHAPIRO and S. ADHYA. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

BOTSTEIN, D. and M. J. MATZ, 1970 A recombination function essential to the growth of bacteriophage P22. J. Mol. Biol. **54**: 417–440.

BOULTER, J. and N. LEE, 1973 Isolation of specialized transducing lambda phages carrying the 1-arabinose operon of *E. coli* B/R. J. Bacteriol. **123**: 1043–1054.

BRENNER, M. and B. N. AMES, 1971 The histidine operon and its regulation. pp. 349–380. In: *Metabolic Regulation*; Volume 5 of *Metabolic Pathways.* Edited by H. J. VOGEL. Academic Press, New York.

BUKHARI, A. I., J. A. SHAPIRO and S. ADHYA, (editors), 1977 *DNA Insertion Elements, Plasmids and Episomes.* Cold Spring Harbor Lab., Cold Spring Harbor, New York.

BUKHARI, A. and D. ZIPSER, 1972 Random insertion of Mu-*1* DNA within a single gene. Nature **236**: 240–243.

CHAN, R. K., 1974 Specialized transduction by bacteriophage P22. Ph.D. Dissertation. M.I.T.

CHAN, R. K., D. BOTSTEIN, T. WATANBE and Y. OGATA, 1972 Specialized transduction of tetracycline resistance by phage P22 in *Salmonella typhimurium.* II. Properties of a high frequency transducing lysate. Virology **50**: 883–898.

CHUMLEY, F., G. FORREST, R. MENZEL and J. R. ROTH, 1979 Hfr formation directed by Tn10. Genetics **91**: 639–655.

DANIELL, E., R. ROBERTS and J. ABELSON, 1972 Mutations in the lactose operon caused by bacteriophage Mu. J. Mol. Biol. **69**: 1–8.

FOSTER, T. J., 1977 Insertion of the tetracycline-resistance translocation unit Tn*10* into the *lac* operon of *E. coli* K12. Molec. Gen. Genet. **154**: 305–309.

HEFFRON, F., P. BEDINGER, J. CHAMPOUX and S. FALKOW, 1977 Deletions affecting transposition of an antibiotic resistance gene. pp. 151–160. In: *DNA Insertion Elements, Plasmids and Episomes.* Edited by A. I. BUKHARI, J. A. SHAPIRO and S. ADHYA. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

HOPPE, I., H. M. JOHNSON, D. BIEK and J. R. ROTH, 1979 A refined map of the *hisG* gene of *Salmonella typhimurium.* Genetics **92**: 17–26.

ISRAEL, V., 1967 The production of inactive phage P22 particles following induction. Virology **33**: 317–322.

KLECKNER, N., 1977 Translocatable elements in prokaryotes. Cell **11**: 11–23.

KLECKNER, N., R. K. CHAN, B. K. TYE and D. BOTSTEIN, 1975 Mutagenesis by insertion of a drug resistance element carrying an inverted repetition. J. Mol. Biol. **97**: 561–575.

KLECKNER, N., J. ROTH and D. BOTSTEIN, 1977 Genetic engineering *in vivo* using translocatable drug resistance elements: New methods in bacterial genetics. J. Mol. Biol. **116**: 125–159.

KRETSCHMER, P. and S. N. COHEN, 1977 Selected translocation of plasmid genes: Frequency and regional specificity of translocation of the Tn*3* element. J. Bacteriol. **130**: 888–899.

LURIA, S. E. and M. DELBRUCK, 1943 Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**: 491–511.

MILLER, J. H., 1972 *Experiments in Molecular Genetics.* Cold Spring Harbor Press, Cold Spring Harbor, New York.

NATHANS, D. and H. O. SMITH, 1975 Restriction endonucleases in the analysis and restructuring of DNA molecules. Ann. Rev. Biochem. **44**: 273–293.

PFEIFFER, D., P. HABERMANN and O. KUBAI-MARONI, 1977 Specific sites for integration of IS elements within the transferase gene of the *gal* operon of *E. coli* K12. pp. 31–36. In: *DNA Insertion Elements, Plasmids, and Episomes.* Edited by A. I. BUKHARI, J. A. SHAPIRO and S. ADHYA. Cold Spring Harbor, New York.

PFEIFER, D., H. J. HIRSCH, D. BERGMANN and M. HAMLAOUI, 1974 Non-random distribution of deletion endpoints in the *gal* operon of *E. coli.* Molec. Gen. Genet. **132**: 203–213.

REUBENS, C., F. HEFFRON and S. FALKOW, 1976 Transposition of a plasmid DNA sequence which mediates ampicillin resistance: independence from host *rec* functions and orientation of insertions. J. Bacteriol. **128**: 425–434.

ROSS, D. G., J. SWAN and N. KLECKNER, 1979 Physical structures of Tn10-promoted deletions and inversions: role of 1400 basepair inverted repetitions. Cell **16**: 721–731.

SCHMEISSNER, U., D. GANEM and J. H. MILLER, 1977 Genetic studies of the *lac* repressor. II. Fine structure deletion map of the *lac*I gene and its correlation with the physical map. J. Mol. Biol. **109**: 303–326.

SCHMIEGER, H., 1972 Phage P22 mutants with increased or decreased transduction abilities. Molec. Gen. Genet. **119**: 75–88.

SHARP, P. A., S. N. COHEN and N. DAVIDSON, 1973 Electron microscope heteroduplex studies of sequence relations among plasmids of *E. coli.* II. Structure of drug resistance (R) factors and F factors. J. Mol. Biol. **75**: 235–255.

SHAMADA, K., R. WEISBERG and M. E. GOTTESMAN, 1973 Mutations induced by bacteriophage lambda in *E. coli* K12. J. Mol. Biol. **80**: 297–314.

SMITH, H. O. and M. LEVINE, 1967 A phage P22 gene controlling integration of prophage. Virology **31**: 207–216.

STARLINGER, P. and H. SAEDLER, 1976 IS Elements in microorganisms. Curr. Top. Microbiol. Immunol. **75**: 111–152.

WEINSTOCK, G., 1977 Genetic and physical studies of bacteriophage P22 genomes containing translocatable drug resistance elements. Ph.D. Dissertation, M.I.T.

WEINSTOCK, G. and D. BOTSTEIN, 1978 Regional specificity of illegitimate recombination associated with the translocatable ampicillin-resistance element Tn*1*. Cold Spring Harbor Symp. Quant. Biol. **43**: 1209–1215.

Corresponding editor: H. ECHOLS