Mapping of *Escherichia coli* Chromosomal Tn*5* and F Insertions by Pulsed Field Gel Electrophoresis

Cassandra L. Smith* and Richard D. Kolodner†

**Departments of Microbiology and Psychiatry, College of Physicians and Surgeons, Columbia University, New York, New York 10032, and †Laboratory of Molecular Genetics, Dana Farber Cancer Institute, Boston, Massachusetts 02115*

Manuscript received September 14, 1987
Revised copy accepted February 22, 1988

ABSTRACT

A low resolution *Not* I physical map of *Escherichia coli* was recently constructed. In this report we demonstrated that this map can be used to map Tn*5* and F insertions physically. The transposon, Tn*5*, contains *Not* I recognition sequences in its IS50 sequences. F plasmid contains an unmapped *Not* I site. Hence, the location of Tn*5* and F in the chromosome can be mapped by identifying the location of the introduced *Not* I sites using pulsed field gel electrophoresis. The physical mapping of genetically mapped Tn*5* insertions confirm the previously constructed *Not* I map and helps align the *E. coli* physical and genetic maps. The use of Tn*5* can assist the construction of both physical and genetic maps for microorganisms lacking such maps. Variations on this approach will facilitate physical mapping with a wide variety of organisms, enzymes, and genetic elements.

THE restriction enzyme *Not* I fragments the *Escherichia coli* chromosome into 22 pieces averaging 200 kb in size (SMITH *et al.* 1986). Recently, these fragments were ordered around the chromosome (SMITH *et al.* 1987c). The existence of this *E. coli* physical map now allows genetic mapping to be done directly by observing alterations in chromosomal DNA structure.

A survey of the DNA sequence bank, Genbank, revealed that the insertion sequence IS50 (AUERSWALD, LUDWIG and SCHALLER 1980) contains a *Not* I recognition sequence. The transposon, Tn*5*, is composed of three antibotic resistance genes bordered by two IS50 elements (MAZODIER *et al.* 1985; BERG and BERG 1987). This means that the location of Tn*5* is marked by *Not* I sites (Figure 1). Physical mapping of the introduced *Not* I sites permits the accurate mapping of Tn*5* insertions. In this report we demonstrate that large DNA technology can be applied to mapping the location of Tn*5* insertions on the *E. coli* chromosome and aid in the genetic analysis of complex phenotypes. The approach also can be used to physically map the location of F DNA in the chromosome.

The data presented here further confirm the *E. coli Not* I map constructed previously. In addition, the physical mapping of the genetically mapped Tn*5* insertions aids in the alignment of the physical and genetic maps. The ease of construction of physical maps combined with the promiscuity of Tn*5* will allow this approach to be applied directly to a wide variety of microorganisms. Thus, the use of Tn*5* can facilitates the efficient construction of both genetic and physical maps. Variations on this approach will allow it to be applied to a variety of organisms, enzymes, and genetic elements.

MATERIALS AND METHODS

E. coli sample preparation, restriction enzyme digestion, and pulsed field gel (PFG) electrophoresis conditions are described elsewhere (SMITH *et al.* 1986; SMITH and CANTOR 1987). The only modification was that the DNA insert samples were prepared directly from overnight L broth cultures. All samples were single colony isolated at the time of their construction and prior to the beginning of these experiments. At that time the samples were grown to stationary phase twice. This ensured that there would be some secondary Tn*5* hops. Blotting, labeling, and hybridization conditions have been described (SMITH and CANTOR 1987; SMITH *et al.* 1987a). Clones used as probes in hybridization experiments are listed in Table 1. The *E. coli* strains used in this study are listed in Table 2. All strains, except SDB1047, are isogenic to AB1157. The location of Tn*5* in six strains that have been previously genetically mapped is shown in Tables 2 and 3.

Three strains contained unmapped Tn*5* insertions—RDK1634, RDK1635, RDK1639. These were derived from JC9604 (GILLEN, WILLIS and CLARK 1981) by first isolating single Tn*5* insertions in JC9604 as described by KOLODNER, FISHEL and HOWARD (1985) and then screening individual mutants to identify those in which the frequency of plasmid recombination was reduced (DOHERTY, MORRISON and KOLODNER 1983) using methods developed by C. A. LUISI-DELUCA and R. D. KOLODNER which will be described elsewhere.

RESULTS

Correlation of the genetic and physical maps of *E. coli* AB1157: The physical map of *E. coli* was constructed with the *E. coli* K12 wild type strain, EMG2 (SMITH *et al.* 1987b). The bacterial strains used in these experiments were almost all derived from

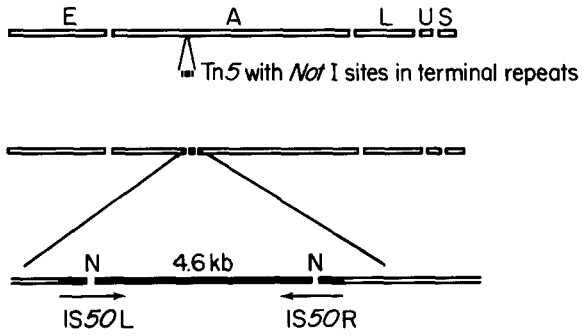

FIGURE 1.—Physical mapping of Tn*5* insertions. The location of Tn*5* is marked by the introduction of two new *Not* I sites located in the IS50 sequences.

TABLE 1

Probes used for hybridization experiments

Probe	Gene	Location (min)[a]	*Not* I fragment[b]	Reference
pEYDG1	Tn*5*			YAKOBSON and GUINEY (1984)
pLC1	*ara*	1	D	CLARK and CARBON (1975)
pUE	*trkE*	29	I	HILL, HENSON and KUEMPEL (1987)
pBS9	(260–310 kb)[c]	32	I	BEJAR and BOUCHE (1983)
colE1-*cdd*	*cdd*	46	F	J. JOSEPHSON and P. VALENTIN-HANSEN (personal communication)
pSE411	*hisS*	53	E	EISENBEIS and PARKER (1982)
pKB45	*pheA*	57	E	ZURAWSKI *et al.* (1978)
pLC1842	*recA*	58	A	SANCAR and RUPP (1979)
pPR42	*tolC*	66	A	HACKETT and REEVES (1983)
pTK201	*glyS*	79	A	WEBSTER *et al.* (1983)
pPP1	*pyrE*	82	L	POULSON *et al.* (1983)

[a] Refers to location on genetic map (BACHMANN 1987a).
[b] See SMITH *et al.* (1987b) and Figure 1.
[c] Refers to location in kilobases on BOUCHE's (1982) 470-kb map of the *E. coli* terminus region.

TABLE 2

***E. coli* strains used in this study**

Strain	Relevant genotype	Reference
AB1157	Wild type[a]	BACHMANN (1987b)
SDB1047	*sbcA*::Tn*5*[b]	FOUTS *et al.* (1983)
JC9604[c]	*recA56, recB21, recC22, sbcA23*(*rac*$^+$)	GILLEN, WILLIS and CLARK (1981)
RDK1372[c,d]	$\Delta^{(recA\text{-}srl)}304^{supB}$	CSONKA and CLARK (1979)
RDK1397[c]	*lexA*::Tn*5*	KOLODNER, FISHEL and HOWARD (1985)
RDK1540[c]	*recN*::Tn*5*	Same as above
RDK1541[c]	*recO*::Tn*5*	Same as above
RDK1516[c]	*uvrD*::Tn*5*	FISHEL, SIEGEL and KOLODNER (1986)
RDK1517[c]	*mutS*::Tn*5*	Same as above
RDK1634[c,e]	*rac*$^-$, ?::Tn*5*	This study
RDK1635[c,e]	*rac*$^-$, ?::Tn*5*	This study
RDK1639[c,e]	*rac*$^-$, ?::Tn*5*	This study

[a] The genotype of AB1157 is *thr-1, leuB6, lacY1, galK2, ara-14, xyl-5, mtl-1, proA62*(*gpt* − *proA*)$_\Delta$; *hisL4, aroE3, str-31, tsx-33, supE44,* and *kdgK51.*
[b] SDB1047 is a Tn*5* induced, mitomycin C resistant derivative of JC5491 (Hfr (PO45), *sbcA*$^+$, *recB21, recC22, thr300, ile318*). Genetic mapping showed the Tn*5* had inserted in the *rac* locus to create an *sbcA* mutation *sbcA119*::Tn*5*.
[c] Isogenic to AB1157.
[d] RDK1372 is a *supB* derivative of JC10283. The *supB* mutation was introduced using a lambda *supB* transducing phage (KATO *et al.* 1980).
[e] Derived from JC9604; see text.

TABLE 3

Expected results of physical mapping experiments of *E. coli* strains containing Tn*5* insertions

Strain	Tn*5* mutation	Min	*Not* I fragments[a] Name	Size(kb)	Expected size of new *Not* I fragments[b]
SDB1047	*sbcA119*	30	I	230	75 + 154
RDK1397	*lexA71*	92	H	230	60 + 170
RDK1516	*uvrD260*	85	S,T	40 or	15 + 25
			L	203	3 + 200
RDK1517	*mutS201*	59	A	1000	50 + 950
RDK1540	*recN1502*	57	E	250	50 + 200
RDK1541	*recO1504*	55	E	250	125 + 125

[a] This fragment should disappear from the *Not* I digest.
[b] This assumes an average colinearity between the genetic and physical map.

AB1157. Since most genetic differences between AB1157 and EMG2 are well characterized, physical mapping studies using AB1157 does not necessarily require that its complete physical map be constructed. However, there is very good correlation between the fragments generated by digeston of AB1157 DNA with *Not* I and its known genetic deviations from EMG2 (Figure 2). PFG separation of *Not* I fragments generated from EMG2, AB1157, and JC9604 (the AB1157 derivative used in this study) is shown in Figure 3. A PFG pulse time of 25 sec resolves most of the *E. coli Not* I fragments. A PFG separation using 100-sec pulse times is required to resolve those fragments above 300 kb. (For example, see Figure 4B.)

EMG2 contains a lambda prophage and an F$^+$ plasmid; AB1157 does not. Therefore, two differences in the *Not* I digestion pattern of the two strains are a shift in fragment K from 205 kb to 160 kb due to the absence of lambda prophage (48.5 kb) from this fragment and the absence of fragment Q which is the F$^+$ plasmid. These changes can easily be seen

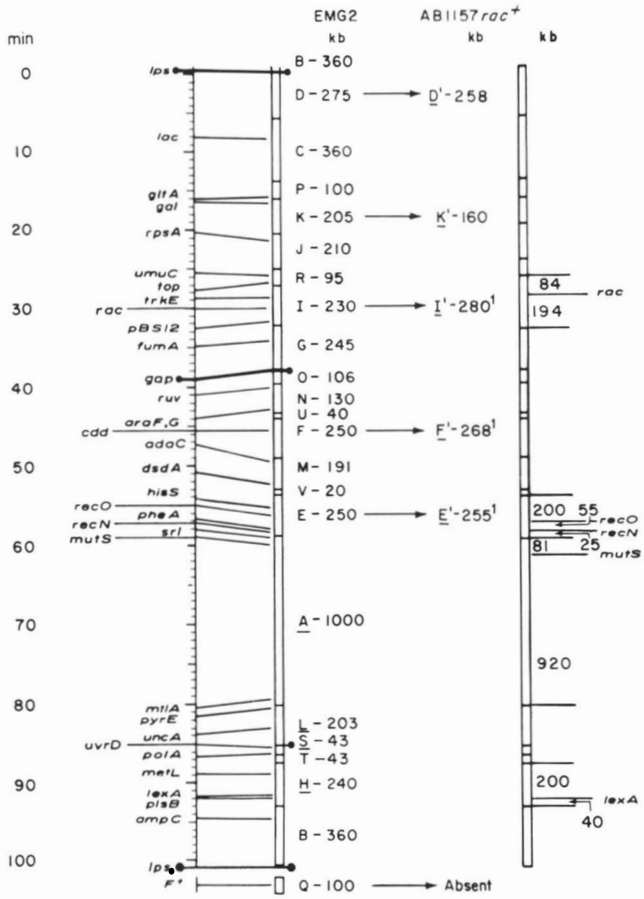

FIGURE 2.—The *Not* I physical map of AB1157. The genetic map is summarized from BACHMANN (1987a). The physical map of EMG2 is summarized from SMITH *et al.* (1987b). Indicated are the observed deviations of AB1157 from EMG2. The sizes of fragments D′, I′, E′ and F′ are taken from the data for strain JC9604, the AB1157 derivative used in this study. Genes not previously mapped by physical methods are shown on the left hand side. Fragments underlined have been confirmed in JC9604 in this study. The alignment of the physical and genetic map is shown in the right hand side. Indicated are the distances (kb) from genes containing Tn*5* insertion to the ends of the parent *Not* I fragments.

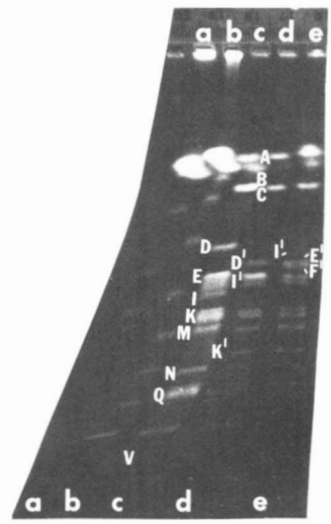

FIGURE 3.—A comparison of the *Not* I fragments of EMG2 and AB1157. The DNA fragments were separated on a 1% agarose gel run for 40 hr at 10 V/cm in a LKB Pulsaphor apparatus using a 25-sec pulse time. Some of *Not* I fragments are identified on the figure. A complete listing may be found in SMITH *et al.* (1987b). Fragments modified in AB1157 and JC9604 are indicated by (′). Lane: a, *S. cerevisiae*; b, concatemerized λcI$_{857}$; c, EMG2; d, AB1157; e, JC9604.

in the photographs of the ethidium bromide stained PFG gels shown in Figures 3, 4, and 5. In addition, AB1157 normally does not contain the *rac* prophage (TEMPLIN, KUSHNER and CLARK 1972). Thus band I in AB1157 should have a lower molecular weight than band I in EMG2. However, the AB1157 derivatives used in this study are derived from JC9604 which contains the *sbcA23* (*rac*$^+$) allele. The hybridization results in this study (see below) show that fragment I is about 200 kb in RDK1184 and 230 kb in EMG2. AB1157 also contains a small deletion, *pro62*(*gpt-proA*)Δ in fragment D. This change is quite apparent when a *Not* I digest of AB1157 and EMG2 are resolved in adjacent lanes during PFG electrophoresis (Figure 3). The specific identification of quite a few of the other *Not* I fragments of AB1157 was confirmed during the hybridization experiments discussed below. Fragment F appears to be very slightly larger in strain AB1157 and its derivatives than it is in EMG2. This physical change does not appear to correlate with the genetic history of AB1157. However, AB1157 has been mutagenized many times during its evolution it is quite possible that an unrecognized rearrangement may have occurred because of these mutagenic treatments.

Mapping F$^+$ DNA in the chromosome: This *E. coli* strain was included in this study because it has a *sbcA*::Tn*5* mutation (see below). A *Not* I map for SDB1047 is not shown because it is not relevant to most this study and very little data is available from this work for its construction. However, an interpretation of the ethidium bromide-stained *Not* I digest of SDB1047 DNA (lane E, Figure 3) is described below as an example of how correlations between physical and genetic data can be constructed and as a brief example of physical mapping of Hfr chromosomal integration sites. The detailed approaches discussed below for Tn*5* mapping can also be applied to mapping the location of F in the chromosome. SDB1047 has no F$^+$ plasmid nor lambda prophage integrated into the chromosome. Thus, expected changes in *Not* I fragments, K and Q, relative to EM62 discussed above are also applicable to this strain. Strain SDB1047 has an F factor integrated around 62 min on the genetic map (LOW 1987). F$^+$ has one unmapped *Not* I site (SMITH *et al.* 1987b). Thus, the introduction of F DNA into the chromosome will lead to the generation of a new *Not* I site if the F$^+$ *Not* I site is retained. The genetic mapping

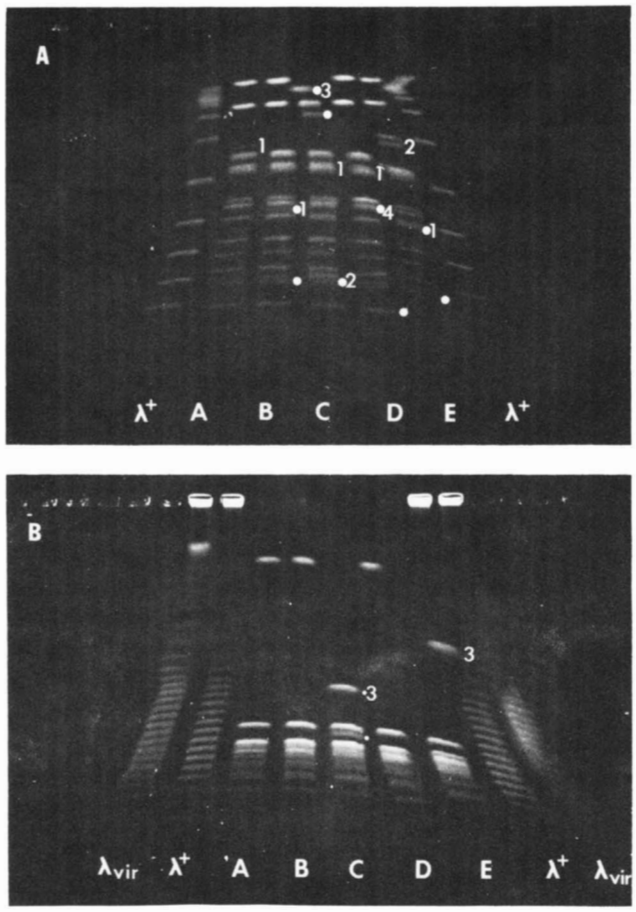

FIGURE 4.—Ethidium bromide-stained PFG electrophoretic separation of *Not* I digested chromosomal DNA from three *E. coli* strains containing unmapped Tn*5* insertions. Experimental details are the same as in Figure 3. (A) Pulse time were 25 sec and (B) 100 sec. Annealed λcI_{857} DNA was used as a size standard on the outside land of (A) and in the next to outside lanes in (B). Annealed λvir DNA (42.5 kb) was also run in (B) in the outside lanes. The *Not* I digested bacterial samples, left to right, are A, JC9604; B, RDK1634; C, RDK1635; D, RDK1639; and E, SDB1047. Probes that hybridized to specific bands were: ○, pEYDG1 (Tn*5*); 1, pBS12; 2, pLC1842; and 3, colEl-*cdd*.

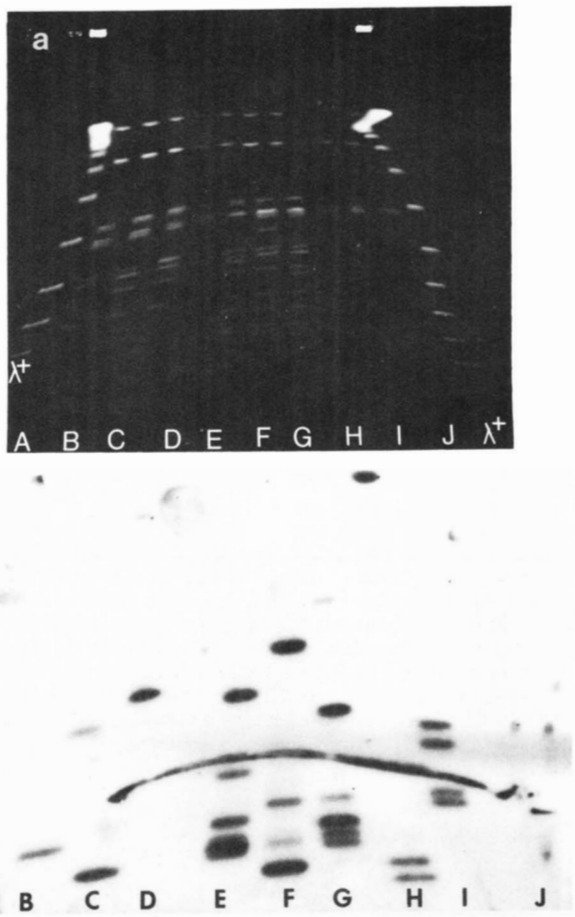

FIGURE 5.—Hybridization with Tn*5* of PFG separated *Not* I fragments from *E. coli* chromosomal DNA. PFG running conditions were the same as in Figure 3a. (a) Ethidium bromide-stained gel and (b) autoradiograph of a Southern blot of the PFG gel hybridized with Tn*5*. The outside lanes contained annealed λ vir DNA size standard. The *Not* I digested bacterial DNA samples are, A, JC9604; B, RDK1634; C, RDK1639; D, RDK1372; E, RDK1397; F, RDK1540; G, RDK1541; H, RDK1516; I, RDK1517; and J, RDK1041.

of Hfr(P045) predict that this integration event is into the *E. coli Not* I fragment A. This means that fragment A should be converted to two new fragments. Fragment A in SDB1047 is diminished in size to 825 kb (lane E, Figure 4B). A new fragment appears at about 255 kb (lane E, Figure 4A). The combined sizes of these fragments (1080 kb) suggests they represent fragment A plus F DNA. The strain also appears to have one or two other changes that involve fragments D, E or F.

The results in this and the previous section show that both AB1157 and SDB1047 contain previously unrecognized chromosomal deviations from EMG2. The recognition of such genomic rearrangement should help clarify genetic mapping data or other experimental data obtained with these strains. The results underscore the utility of PFG characterization of even genetically well characterized strains.

Approaches to physical mapping of Tn*5* insertions: Since IS50 sequences contain recognition sequences for the restriction enzyme *Not* I, the location of Tn*5* is marked by such sites (Figure 1). The *Not* I recognition sequence is 8 bp in size and occurs only 22 times in the entire *E. coli* chromosome (SMITH *et al.* 1987b). Thus, the simplest way to accurately map *E. coli* chromosomal Tn*5* insertions is to identify the newly introduced *Not* I sites in chromosomal DNA. This can be accomplished by digesting chromosomal DNA and analyzing the fragments by PFG electrophoresis. Ethidium bromide staining of such gels should reveal differences between isogenic strains that are due to the Tn*5* insertion (Figures 4 and 5). If this is the only approach being used, the strains being compared should be isogenic, and the physical

TABLE 4

Summary of PFG analysis of mapped and unmapped Tn*5* insertions

Strain	Deleted fragments[1]	No. of insertions[2]	Fragments which hybridize to Tn*5*[2]	*Not* I fragment[3]
SDB1047	—	1	87[a] + 194[b] = 281	I*
		2-vl		
RDK1397	H	1	40 + 200 = 240	H
		2-M	30 + 45 + 64 + 193	F*
RDK1516	S-T	1	15 + 20 = 35	S or T
RDK1517	D, H	3	73 + 81[c] + 134 + 153 + 757[d]	A*, D*
RDK1540	—	1	25 + 226[e] = 251	E*
RDK1541	—	1	80 + 176[e] = 255	E*
		2-m	280 + 81 + 55 + 42	B or C and O, P, or R
RDK1634	—	1	85 + 200[b] = 285	I*
RDK1635	—	2	97[f] + 330[d] + 618[g] = 1045	A*
RDK1639	—	1	45 + 220[h] = 265	F*

[1] These bands were obviously missing from the ethidium bromide stained gel. In some cases, indicated by —, it was not possible to see the deleted band. This is probably due to the fact that in these experiments the chromosome replication fork were not synchronized. Thus the yield of each fragment depended on its distance from the replication origin (see RESULTS).

[2] Major bands are listed in the first row; minor bands are listed in the second row for each century. The intensity of the hybridization signal with Tn*5* for the minor bands is indicated vl = very light, m = medium. Fragments indicated hybridized to the following probes: *a* = pUE, *b* = pBS9, *c* = pLC1, *d* = pPR42, *e* = pSE411, *f* = pLC1842, *g* = pTK201, and *h* = colE1-*cdd*.

[3] The location was determined by a combination of analyzing missing parental *Not* I fragments, size determination of the IS50-genomic fragments, and/or hybridization experiments with putative nearby chromosomal genes. Those fragments that were confirmed by hybridization with nearby cloned chromosomal genes are indicated with an *.

map should be known for the particular pair of strains.

The Tn*5* containing strains will be missing one previously seen *Not* I fragment that has now been split into three new *Not* I fragments (Tables 3 and 4). One new fragment, containing the antibiotic resistant genes of the Tn*5* element, will be very small, 4.6 kb. The amount of this fragment may be too low to be seen by ethidium bromide staining, and in fact it may run off the bottom of the gel under some PFG conditions. The other two new *Not* I fragments represent the IS50-genomic junction fragments. Their combined size should roughly equal the size of the fragment that has disappared from the *Not* I digest. The identification of the parent *Not* I fragment will approximately locate the insertion. The sizes of the new *Not* I fragments will reveal the distance from the end of the *Not* I fragment to the Tn*5* insertion site. Since there are two possible orientations of the new *Not* I fragments relative to the ends of the parent *Not* I band, two possible integration sites will be identified. Hybridization experiments with Tn*5* can be used to confirm that the shifts seen in the ethiduim bromide stained patterns are due to Tn*5* insertions (Figures 4 and 5). Three fragments should hybridize to Tn*5*: the internal 4.6-kb fragment and the two IS50-genomic fragments. Hybridization experiments with genomic clones will clarify ambiguities that could arise from several fragments having the same size, confirm the putative physical map of strains for which the complete map has not been constructed, and, most important, precisely locate the Tn*5* element.

Precise mapping requires that the orientation of the new *Not* I fragments be determined. This is done by using a clone from the putative split fragment as a probe in a hybridization experiment to determine the distance of the Tn*5* insertion from an end of the parental *Not* I fragment (Table 4 and Figure 5). Using this approach we have physical mapped six genetically mapped Tn*5* insertions and three unmapped Tn*5* insertions.

Physical mapping of Tn*5* insertions correlates the genetic and physical maps: The expected results from physically mapping the genetically mapped Tn*5* insertions are indicated in Table 3. The hybridization results with Tn*5* are shown in Figures 5 and are summarized in Table 4 and Figure 2. Hybridization with Tn*5* detects three fragments, a 4.6-kb internal fragments containing the antibiotic resistant genes, and the two IS50-genomic junction fragments. The results show that there is very good correlation between the genetic map and the physical map when the *Not* I fragment is identified as the sum of the genomic-IS50 junction fragments. For instance, the junction fragments in strain RDK1397 are 40 kb and 200 kb. This means that the insertion is into a fragment that is 240 kb. The physical map of EMG2 predicts that this insertion is into band H whose size was previously determined to be 230 kb. Fragment H is missing in the ethidium bromide-stained gel of

Not I digested RDK 1397 which further confirm this interpretation.

The locations of the *sbcA*, *lexA*, *uvrD*, *mutS*, *recN*, and *recO* genes have been characterized by P1 transduction. In all but one of these cases, the observed genomic-IS50 fragments were within 25 kb of the predicted size based on the physical map. This may, in part, reflect the precision of the genetic mapping of these genes, and the alignment of the physical map on the genetic map. The greatest deviation between the physical map and the genetic map was obtained with the *recO* gene. This gene was 45 kb from its predicted location. This gene was mapped by 2- and 3-factor crosses (KOLODNER, FISHEL and HOWARD 1985). Cloning experiments showed that this gene was within 10 kb of the *lep* gene (P. MORRISON, S. T. LOVETT, L. GILSON and R. D. KOLODNER, unpublished results) which is consistent with its genetic location. These results suggest that there is a distortion in the linkage distances in this region of the genetic map. Fragment E appears to be larger in JC9604 than in EMG2 (Figure 2); thus such a distortion may be limited to JC9604.

The expected sizes of the new *Not* I fragments summarized in Table 3 were obtained by assuming an average of 46.5 kb per min. The actual variation needed to align the *E. coli* EMG2 physical and genetic map (SMITH *et al.* 1987b) was between 61 kb and 38 kb. In most cases only the approximate location of the ends of the *Not* I fragments is known. A precise physical map and its accurate alignment with the genetic map requires that the *Not* I sites be located both genetically and physically. In fact, the results of these experiments with Tn*5* help identify the exact location of the natural *Not* I sites in the chromosomes. For instance, in RDK1397 the *lexA* gene is 40 kb from the end of the distal *Not* I site. Figure 2 summarizes the alignment data obtained from these studies.

These experiments also allow one to determine the distance between genes. For instance, strains RDK1540 and RDK1541 both have insertions in fragment E. Since plasmid pSE411 hybridized to the new fragments in the two strains, 226 kb and new 178 kb respectively (Table 4), the distance between the genes is simply the different in fragment size, 48 kb.

Direct physical mapping of unmapped Tn*5* insertions: Three of the *E. coli* strains studied, RDK1634, RDK1635, and RDK1639, contained unmapped Tn*5* insertions. A comparison of the ethidium bromide patterns obtained from these three strains following a *Not* I digest with JC9604 revealed some differences (Figures 4 and 5). The most obvious difference (lane C Figure 4) is the splitting of fragment A in strain RDK1635 into two smaller fragments. (Hybridization experiments described below will show that fragment A was actually split into three new fragments.) Strain RDK1639 is missing a fragment from the E-F doublet and one new fragment appears coincidental with fragment I at 210 kb. Strain RDK1634 does not appear to be missing any fragments but a new very light fragment has appeared at 200 kb, between fragments M and L.

Hybridization results using pEYDG1(Tn*5*) as a probe (Figure 4 and 5 and Table 4) prove the changes observed on ethidium bromide stained gels are due to Tn*5* insertions and locate changes due to Tn*5* insertions that cannot be seen in the ethidium bromide stained gels. The exact location of the transposable element was proven by hybridization experiments with putative nearby gene specific probes as described below.

RDK1634: *E. coli* strain RDK1634 has Tn*5* integrated into the *rac* prophage located on fragment I. This was shown by several criteria. The sum of the IS50-genomic fragments for this strain was 285 kb. The PFG gel shown in Figure 3 was hybridized to pBS12, a probe from one end of fragment I that did not include the *sbcA* locus. The results showed that fragment I was 281 kb in the parent strain JC9604 and that this probe detected only a 200 kb fragment in RDK1634.

Fragment I was similar in size to the fragment detected in SDB1047 by hybridization with pBS12. SDB1047 was included in this study, although it is not isogenic to the other strains, because it has a geneticaly mapped *sbcA*::Tn*5* mutation. The results indicate that in SDB1047 fragment I is about 274 kb rather than the 230 kb observed in wild type. This is very slightly smaller than the fragment I in the JC9604 derivatives. This may be due to the fact that SDB1047 has a different *sbcA* allele than JC9604.

The fragments detected by pBS12 stained very lightly with ethidium bromide. This result supports the notion that the fragment was from the replication termination region which should be underepresent stoichiometrically. The samples were prepared under conditions where there should be multiple replication forks present in each cell. Therefore, the yield of *Not* I fragments from regions of the chromosome close to the replication origin could be two- to fourfold greater than those from the replication termination region. The bias yield can be avoided, if desired, by synchronizing replication by incubating the cells in chloramphenicol for one hour before extracting the DNA (SMITH *et al.* 1986; SMITH and CANTOR 1987). This treatment allows the ongoing round of replication to finish but does not allows DNA initiation. The EMG2 DNA sample used in lane C Figure 3 was the only sample in the present study extracted from synchronized cells. Note how much darker fragment

I is in EMG2 than fragment I′ in the other *E. coli* strains. Thus, one should be cautious in interpreting on ethidium bromide-strained results obtained using DNA samples from unsynchronized cells. (See the Band Missing data in Table 4.) Synchronization is not as critical for hybridization experiments as shown by the results presented here.

RDK1635: *E. coli* strain RDK1635 appears to have two Tn*5* insertions in fragment A. This was revealed when hybridization experiments with Tn*5* showed equal signals from three bands (97, 330 and 618 kb). The fact that the hybridization signal detected by pEYDG1 (Tn*5*) was equal from the three bands suggests that RDK1635 had multiple insertions present in the entire population of cells. These results were obtained despite the fact that the recipient cells were single colony isolated several times immediately after the hopping experiment and before DNA samples were made. These results certainly underscore the need to check for multiple insertions and/or early transposition events that could spread through the entire population (see below). This strain also appears to have a deletion of the *rac* allele (see below).

The three IS50-genomic junction fragments derived from fragment A were subsequently further characterized to locate the insertions exactly. Plasmid pLC1842, containing the *recA* gene at 58 min, hybridized to the 97-kb fragment while pTK201, containing the *glyS* at 79 min, hybridized to the 618-kb fragment. In addition, hybridization experiments with pLC1842 to a partial digest of RDK1635 indicated that a 330-kb fragment was adjacent to the 97-kb fragment (data not shown). These results show that the Tn*5* insertions are at 60 min and 68 min.

RDK1639: *E. coli* strain RDK1639 has a single Tn*5* insertion in fragment F. *Not* I fragments E and F are the same size. They were distinguished by hybridization experiments using pKB45 to detect fragment E and plasmid colE1-*cdd* to detect fragment F. The results indicate that the insertion is in fragment F around 46 or 47 min.

The Rec$^-$ phenotype of RDK1635 and RDK1639 may be due to deletion of the *rac* allele: In addition to the Tn*5* insertions, RDK1635 and RDK1639 have a 25-kb deletion in fragment I which was detected by hybridization with pBS12 (Figure 4 and Table 4). Since *rac* prophage is on this fragment, and the selection procedure would have picked up *rac*$^-$ mutations, it is possible that the phenotype of these strains is due to a *rac* deletion rather than a Tn*5* insertion. Furthermore, the Rec$^-$ phenotype did not cotransduce with Tn*5*. *Rac*$^-$ mutations can be deletions, duplications or point mutations of the *sbcA* locus (KAISER and MURRAY 1980). RDK1635 and RDK1639 both appear to contain deletions of this locus. It is not clear whether there is a relationship between the Tn*5* insertions and the *rac* deletions. These could be independent events, since mutants were detected at very low levels or, in fact, the Tn*5* induced mutations could promote the *rac* deletions.

Secondary sites of Tn*5* integration: In contrast to genetic studies which involve the characterization of single cells, the experiments presented here examine populations of cells. This means that most Tn*5* hops will be visible, despite the fact that they are present in only a subset of the population. The primary integration site can be distinguished from secondary sites by examining the intensity of the hybridization signal obtained with Tn*5* (Table 4). The primary site will give the darkest signal in hybridization experiments using Tn*5* as a probe since this site is present in the entire population. Likewise it should be represented by an obvious change in an ethidium bromide-stained gel photograph. The experiments outlined in this paper was designed to allow one not only to assess the primary location of Tn*5* but also its spread to other locations in the genome. Hybridization experiments with pEYDG1 (Tn*5*) detected only single insertions in just four of the Tn*5* containing strains (RDK1634, RDK1639, RDK1516, and RDK1540). This was shown by the fact that only two IS50-genomic junction fragments were detected in each of these strains.

As discussed above, *E. coli* strain RDK1635 has two Tn*5* insertions in fragment A since the signal obtained with Tn*5* appears to be equal for each of the three fragments. *E. coli* strain RDK1517 has five fragments that hybridize with what appears to be equal intensity to Tn*5*. This indicates that there are three Tn*5* insertions in RDK1517. The three remaining *E. coli* strains studied appear to have a primary insertion site and a variable number of secondary sites which may be present in at least half the population as judged by the weaker intensity of the hybridization signal with Tn*5*. Strain RDK1397 shows two high intensity fragments and four medium intensity fragments. Strain RDK1541 has two dark fragments and four medium intense fragments. These results indicate these strains contain a primary insertion site and two secondary sites. It is not clear from these results whether the multiple secondary insertions are actually present in the same cell.

Our results underscore the need to check strains for multiple insertions. In fact, integration of Tn*5* into secondary sites may even be selected for in slow growing mutant strains, where the secondary integration site would enhance growth or survival or subsequently permit excision of a Tn*5* that caused the deleterious effect. The methods described here will be useful in experiments designed to explore such questions.

Now that a physical may of the entire *E. coli*

chromosome is known, it is possible to ask where secondary insertions occur. For instance, some hypothetical transposition mechanisms might favor transposition events to nearby locations. Nearby could be defined by primary or tertiery DNA structure. The experiments reported here allow one to begin to determine where the secondary sites are located. For instance, in *E. coli* strain RDK1635 the insertions were confined to fragment A. Hence they were within 1 megabase (Mb) of each other. Strain RDK1517 has insertions that appear to be in fragments bands A, D, and H. Bands D and H are absent from the ethidium bromide photograph, and shifts in their sizes and in band A were confirmed by hybridization. Thus, in this second case the multiple insertions were confined to one quadrant of the chromosome. In these two examples the insertions appear to be relatively close to each other.

In two other cases, RDK1397 and RDK1541, the sites appear to be widely dispersed around the chromosome, which may or may not reflect tertiery structure. An interpretation of the insertions in the former strain places them in bands, H, F, and R (H and F are confirmed); in the latter case the insertions appear to be in E, B or C, and O, P, or R. This is much too small a sample size to draw firm conclusions, but more extensive experiments, now in progress, should allow us to explore such questions.

CONCLUSIONS

The results in this paper show that the location of Tn*5* insertions in the chromosome may be revealed by simple direct physical mapping experiments. In fact, the construction of physical maps can be aided by the use of Tn*5*, and some of the results presented in this report were used to help construct and confirm the physical *Not* I map of *E. coli*. Certainly Tn*5* can help align physical and genetic maps and its use will speed up physical mapping for organisms whether genetic maps exist or not.

The transposon, Tn*5*, has been used successfully for genetic mapping experiments in an ever increasing number of microorganisms (BRUJIN and LUPSKI 1984). For these organisms it will be extremely simple to construct a physical map by using the very simple partial mapping strategy outlined below. Chromosomal DNAs from 5 to 10 isogenic derivatives containing independent Tn*5* insertions can be partially digested with *Not* I and the fragments separated by PFG electrophoresis. Hybridization of such a gel with Tn*5* will detect a ladder of fragments that extend in both direction around the chromosome. [See Figures 6 and 7 in SMITH *et al.* (1987b) for examples of partial mapping experiments.] Such a ladder can be interpreted to locate the *Not* I sites despite the fact that fragments are being detected in both directions. The resulting partial maps should overlap at some point once enough data is accumulated. The beauty of this approach is that it requires only one probe and that the Tn*5* insertions need not even be genetically mapped. Obviously genetic probes and genetic maps will greatly aid this endeavor but they are not required. Such experiments are now in progress for Legionella (A. MARRA, H. SHUMAN and C. L. SMITH unpublished data). The approaches described here can also be applied to eukaryotic organisms. For instance, in yeast, site-directed integration of cloned genes (ORR-WEAVER, SZOSTAK and ROTHSTEIN 1981) would allow the insertion of rare cutting sites into yeast chromosomes. In mammalian cells mapping of transformed cells lines should be aided by engineering rare cutting sites into retroviruses or to other transforming vectors.

Variations on the basic approach used in *E. coli* may be necessary for some organisms. In those organisms with a high GC content, the number of fragments produced by *Not* I (or *Sfi* I) may be too numerous to order. For instance, with *Mycobacterium* species the chromosomal fragments produced by both a *Not* I and a *Sfi* I digestion average less than 50 kb, whereas megabase fragments are produced by digestion with *Dra* I (C. L. SMITH, W. JACOBS and B. BLOOM, unpublished results) which recognizes the 6-bp sequence, TTTAAA. Thus, a mobile element that functioned in Mycobacterium and contained a *Dra* I site would have to be found or constructed to replaced Tn*5* for the experiments discussed above. However, it should be apparent that there are many options for constructing physical maps efficiently. Many prokaryotic and eukaryotic transposable elements have been identified (BERG and BERG 1981; SHAPIRO 1983). Some of these may have rare cutting sites or such sites can be engineered onto the appropriate elements. Thus this approach can be applied to a wide variety of organisms.

Since the transposition of Tn*5* into an essential gene is a lethal event the above approach cannot be applied directly to mapping of these kinds of genes. However, it should be possible to map such genes by identifying a nearby Tn*5* (or *Not* I site). For instance, a mutation could be linked to a newly made Tn*5* insertion by mutagenizing the mutant strain of interest with Tn*5* and selecting for cotransduction of the mutation of interest with Tn*5*. Chromosomal *Not* I sites can also be generated for mapping of cloned genes by recombining such genes, on appropriate vectors containing rare cutting sites, into the chromosome. For instance, an essential gene could be cloned onto a vector with a rare cutting site. Such a clone could be introduced into a *polA*$^-$ cell (GREENER and HILL 1980), grown in low phosphate medium (ROEDER and COLLMER 1985), or linearized and introduced into a *recB*$^-$*C*$^-$ *sbc*$^-$ cell (BASSETT and

KUSHNER 1984) in order to promote its integration into the chromosome. Selection for an antibiotic resistance present on the plasmid would result in cells that have the plasmid introduced into the chromosome, and in this case their location would be marked by the rare cutting site. Such experiments are now in progress (C. L. SMITH and S. CHANG, unpublished results). Thus, direct physical mapping should also be applicable to the mapping of essential genes.

The authors would like to thank CHARLES R. CANTOR for his support and encouragement and JASON ECONOME, STEPHANIE KLCO and ANDREW SCHUTT for excellent technical assistance. The authors are grateful for support from National Institute of Health (CA39782 and GM26017), Department of Energy (DE-FG02-87ER-GD582), and LKB Produkter, AB.

LITERATURE CITED

AUERSWALD, E. A., G. LUDWIG and H. SCHALLER, 1980 Structural analysis of Tn*5*. Cold Spring Harbor Symp. Quant. Biol. **45:** 107–113.

BACHMANN, B. J. 1987a Linkage map of *E. coli* K12. pp. 807–876. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Edited by C. NEIDHARDT, J. INGRAHAM, K. B. LOW, B. MAGASANIK, M. SCHAECHTER and H. E. UNBARGER. American Society for Microbiology, Washington, D.C.

BACHMANN, B. J. 1987b Deviations and genotypes of some mutant derivatives of *E. coli* K12. pp. 1190–1219. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Edited by C. NEIDHARDT, J. INGRAHAM, K. B. LOW, B. MAGASANIK, M. SCHAECHTER and H. E. UNBARGER. American Society for Microbiology, Washington, D.C.

BASSETT, C. L., and S. R. KUSHNER, 1984 Exonucleases I, III, and V are required for stability of colE1-related plasmids in *E. coli*. J. Bacteriol. **164:** 51–56.

BEJAR, S., and J. P. BOUCHE, 1983 Molecular cloning of the region of the terminus of *E. coli* K12 DNA replication. J. Bacteriol. **153:** 605–609.

BERG, C. M., and D. E. BERG, 1981 Bacterial transposons pp. 107–116. In: *Escherichia coli and Salmonella typhimurium Microbilogy-1981*, Edited by D. SCHLESSINGER. American Society for Microbiology, Washington, D.C.

BERG, C. M., and D. E. BERG, 1987 Use of transposable elements and maps of known insertions. pp. 1071–1109. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Edited by C. NEIDHARDT, J. INGRAHAM, K. B. LOW, B. MAGASANIK, M. SCHAECHTER and H. E. UNBARGER. American Society for Microbiology, Washington, D.C.

BOUCHE, J. P., 1982 Physical map of a 470 × 10^3 base pair region flanking the terminus of DNA replication in the *E. coli* K12 genome. J. Mol. Biol. **154:** 1–20.

BRUJIN, F. J. DE, and J. R. LUPSKI, 1984 The use of transposon Tn*5* mutagenesis in the rapid generation of correlated physical and genetics maps of DNA segments cloned into multicopy plasmids—a review. Gene **27:** 131–149.

CLARK L., and J. CARBON, 1975 Biochemical construction and selection of hybrid plasmids containing specific segments of the *E. coli* genome. Proc. Natl. Acad. Sci. USA **72:** 4361–4365.

CSONKA, L. N., and A. J. CLARK, 1979 Deletions generated by the transposon Tn*10* in the *srl-recA* region in the *E. coli* K12 chromosome. Genetics **93:** 329–343.

DOHERTY, M. J., P. T. MORRISON and R. KOLODNER, 1983 Genetic recombination of bacterial plasmid DNA: physical and genetic analysis of the products of plasmid recombination. J. Mol. Biol. **167:** 539–560.

EISENBEIS, S. C., and J. PARKER, 1982 The nucleotide sequence of the promoter region of *hisS*, the structural gene for histidyl-tRNA synthetase. Gene **18:** 107–116.

FISHEL, R. A., E. C. SIEGEL and R. KOLODNER, 1986 Gene conversion in *E. coli*. J. Mol. Biol. **188:** 146–157.

FOUTS, K. E., T. WASIE-GILBERT, D. K. WILLIS, A. J. CLARK and S. D. BARBOUR, 1983 Genetic analysis of transpon-induced mutations of the *rac* prophage in *E. coli* K12 which affect expression and function of *recE*. J. Bacteriol. **156:** 718–726.

GILLEN, J. R., D. K. WILLIS and A. J. CLARK, 1981 Genetic analysis of the *RecE* pathway of genetic recombination in *E. coli* K12. J. Bacteriol. **145:** 521–532.

GREENER, A., and C. W. HILL, 1980 Identification of a novel genetic element *E. coli* K12. J. Bacteriol. **144:** 312–321.

HACKETT, J., and P. REEVES, 1983 Primary structure of the *tolC* gene that codes for an outer membrane protein of *E. coli* K12. Nucleic Acid Res. **11:** 6487–6495.

HILL, T. M., J. M. HENSON and P. L. KUEMPEL, 1987 The terminus region of the *E. coli* chromosome contains two separate loci that exhibit polar inhibition of replication. Proc. Natl. Acad. Sci. USA **84:** 1754.

KAISER, K., and N. E. MURRAY, 1980 On the nature of *sbcA* mutations in *E. coli* K12. Mol. Gen. Genet **179:** 555–563.

KATO, T., Y. SHINOURA, A. TEMPLIN and A. J. CLARK, 1980 Analysis of UV light induced suppressor mutations in a strain of *E. coli* K12 AB1157; implication for molecular mechanisms of UV mutagenesis. Mol. Gen. Genet. **180:** 283–291.

KOLODNER, R. D., R. A. FISHEL and M. HOWARD, 1985 Genetic recombination of bacterial plasmid DNA: effect of *recF* pathway mututations on plasmid recombination in *E. coli*. J. Bacteriol. **153:** 1050–1056.

LOW, K., 1987 Hfr strains of *E. coli*. pp. 1071–1109. In: *Escherichia coli and Salmonella typhimuriumm Cellular and Molecular Biology*, Edited by C. NEIDHARDT, J. INGRAHAM, K. B. LOW, B. MAGASANIK, M. SCHAECHTER and H. E. UNBARGER. American Society for Microbiology, Washington, D.C.

MAZODIER, P., P. COSSART, E. GIRAUD and F. GASSER, 1985 Complete of the nucleotide sequence of the central region of Tn*5* confirms the presence of three resistance genes. Nucleic Acid Res. **13:** 195–205.

ORR-WEAVER, T. L., J. W. SZOSTAK and R. J. ROTHSTEIN, 1981 Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA **78:** 6354–6358.

POULSEN, P., F. BONEKAMP, K. F. JENSEN and W. S. REZNIKOFF, 1983 Tn*5* transposition and its regulation. Cell **31:** 307–308.

ROEDER, D. L., and A. COLLMEER 1985 Marker exchange mutagenesis of pectate lyase isozyme gene in *Erwinia chrysanthemi*. J. Bacteriol. **157:** 661–664.

SANCAR, A., and W. D. RUPP, 1979 Physical map of the *recA* gene. Proc. Natl. Acad. Sci. USA **76:** 3144–3148.

SHAPIRO, J. A., 1983 pp. 1–608. In: *Mobile Genetic Elements*. Academic Press, New York.

SMITH, C. L., and C. R. CANTOR, 1987 Purification, specific fragmentation, and separation of large DNA molecules. Methods Enzymol. **155:** 449–467.

SMITH, C. L., P. E. WARBURTON, A. GAAL and C. R. CANTOR, 1986 Analysis of genome organization and rearrangements by pulsed field gradient gel electrophoresis. Genet. Engrg. **8:** 45–70.

SMITH, C. L., S. K. LAWRANCE, G. A. GILLESPIE, C. R. CANTOR, S. M. WEISSMAN and F. S. COLLINS, 1987a Strategies for mapping and cloning macro-regions of mammalian genomes. Methods Enzymol. **151:** 461–489.

SMITH, C. L., J. ENONOME, A. SCHUTT, S. KLCO and C. R. CANTOR, 1987b A physical map of the *E. coli* genome. Science **236:** 1448–1453.

TEMPLIN, A., S. R. KUSHNER and A. J. CLARK, 1972 Genetic

analysis of mutations indirectly suppressing *recB* and *recC* mutations. Genetics **72:** 205–215.

WEBSTER, T. A., B. W. GIBSON, T. KENG, K. BIEMAN and P. SCHIMMEL, 1983 Primary structure of both subunits of *E. coli* glycyl *t*-RNA synthetase J. Biol. Chem. **258:** 10637–10643.

YAKOBSON, E. A., and D. G. GUINEY, 1984 Conjugal transfer of bacterial chromosomes mediated by the RK2 plasmid transfer origin cloned into transposon Tn*5*. J. Bacteriol. **160:** 451–453.

ZURAWSKI, G., K. BROWN, D. KILLINGLY and C. YANOFSKY, 1978 Nucleotide sequence of the leader region of the phenylalanine operon of *E. coli*. Proc. Natl. Acad. Sci. USA **75:** 4271–4275.

Communicating editor: D. BOTSTEIN