Copyright © 1987 by the Genetics Society of America

A Tn*10*-*lacZ*-*kanR*-*URA3* Gene Fusion Transposon for Insertion Mutagenesis and Fusion Analysis of Yeast and Bacterial Genes

Olivier Huisman,*,[1] Wendy Raymond,* Kai-Uwe Froehlich,† Patrick Errada,* Nancy Kleckner,* David Botstein† and M. Andrew Hoyt†

**Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, and* †*Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

Manuscript received December 23, 1986
Accepted March 2, 1987

ABSTRACT

We describe here a new variant of transposon Tn*10* especially adapted for transposon analysis of cloned yeast genes; it can equally well be used for analysis of prokaryotic genes. We have applied this element to analysis of the *LEU2*, *RAD50*, and *CDC48* genes of *Saccharomyces cerevisiae*. This transposon, nicknamed mini-Tn*10*-LUK, contains a *lacZ* gene without efficient transcription or translation start signals, an intact *URA3* gene, and a *kanR* determinant. The *lacZ* gene can be activated by appropriate insertion of the element into an actively expressed gene. Other yeast genes can easily be substituted for *URA3* in the available constructs. The mini-Tn*10*-LUK system has several important advantages. (1) Transposition events occur in *Escherichia coli* at high frequency and into many different sites in yeast DNA. It is easy to obtain enough insertions to sensitively define the functional limits of a gene. (2) Transposon insertions can be obtained in a single step by standard transposon procedures and can be screened immediately for phenotype either in yeast or in *E. coli*. (3) The LacZ phenotypes of the insertion mutations provide a good circumstantial indication of the orientation of the target gene. (4) Under favorable circumstances, usable *lacZ* protein fusions are created. (5) Transposon insertion mutations obtained by this method directly facilitate additional genetic, functional, physical and DNA sequence analysis of the gene or region of interest.

THE generation of mutations by transposon insertion can be a powerful analytical technique. We recently described a set of derivatives of the bacterial transposon Tn*10* that are useful for defining the functional limits and regulatory sites of bacterial genes (Way *et al.* 1984). We describe here methods that allow this type of analysis to be extended to regions of the genome of the yeast *Saccharomyces cerevisiae*. These methods provide an alternative to a recently developed Tn*3*-based system (Seifert *et al.* 1986). The work described here involves mutagenesis of yeast genes cloned into *Escherichia coli* plasmids. Vectors appropriate for mutagenesis of genes cloned into phage λ vectors are also available, and specific methods for such mutagenesis are described elsewhere (Snyder, Elledge and Davis 1986).

The transposon vehicles described here can also be used with bacterial target genes for transposon mutagenesis and for *lacZ* fusion analysis of bacterial genes.

All of the phage and plasmid vehicles described here can be obtained from N.K. upon request.

MATERIALS AND METHODS

Media: Minimal growth medium for bacteria was M9, rich medium was LB (Miller 1972), and medium for experiments involving λ was λym (Way *et al.* 1984). Where noted, the antibiotics kanamycin, ampicillin and tetracycline were added to concentrations of 50, 100 and 20 μg/ml, respectively. For detection of β-galactosidase activity, 5-bromo-4-chloro-3-indolyl-βD-galactoside (Xgal) was added to 40 μg/ml. Yeast growth media were SD (minimal) and YEPD (rich) (Sherman, Fink and Hicks 1983). For detection of β-galactosidase activity in yeast, buffered minimal medium (Clifton, Weinstock and Fraenkel 1978) with 40 μg/ml Xgal was used.

Construction of mini-Tn*10*-*lacZ*-*URA3*-*kanR*: The progenitor of mini-Tn*10*-LUK is a derivative lacking the *URA3* gene, mini-Tn*10*-LK. Mini-Tn*10*-LK consists of the two termini of Tn*10*, from *Bcl*I sites outward, flanking the *E. coli lacZ* gene missing codons one through seven and the *kanR* gene from Tn*903* (see Figure 1). Construction of mini-Tn*10*-LK is described in detail by Huisman and Kleckner (1987). The resulting plasmid, pNK1207, is a pBR322 derivative containing mini-Tn*10*-LK inserted in a segment of the *Salmonella* histidine operon. To construct mini-Tn*10*-LUK, the unique *Bam*HI site at the junction between *lacZ* and *kanR* DNA in pNK1207 (Figure 1) was opened and a 1.1-kb *Bgl*II fragment of DNA containing the yeast *URA3* gene was inserted (Rose, Grisafi and Botstein 1984) (the original *Hind*III ends of this *URA3* fragment were converted to *Bgl*II ends with linkers). The resulting plasmid is pMA1023. DNA sequence information for all of the nucleotides in mini-Tn*10*-LUK has been published (Rose, Grisafi and Botstein 1984; Halling *et al.* 1982; Kalnins *et al.* 1983; Oka, Sugisaki and Takanami 1981). Note that mini-Tn*10*-LK transposons containing yeast markers other than *URA3* can be easily constructed by inserting other genes at the unique *Bam*HI site of pNK1207.

[1] Present address: Departement de Biotechnologie, Institut Pasteur, 29 rue du Docteur Roux, 75015 Paris, France.

For transposon mutagenesis of plasmid-borne segments, the mini-Tn*10* transposon insertions on pNK1207 and pMA1023 were crossed onto bacteriophage λ derivatives containing Salmonella *his* operon DNA and other appropriate markers (λNK780: λgt7-*his c1857 Pam80: nin5*) as described by FOSTER *et al.* 1981. For the mini-Tn*10*-LK element, λNK780 was grown on a host containing pNK1207 and a recombinant phage carrying the mini-Tn*10* insertion within the *his* sequences was selected by its ability to confer kanamycin resistance upon a λ lysogen by integration into the resident prophage. The resulting double lysogens were induced to produce phage and the desired λgt7-*his*::Tn*10*-LK phage, named λNK1205, was identified, purified and grown. An isogenic phage containing *his*::Tn*10*-LUK, λNK1224, was made by an analogous procedure; λNK1205 was grown on a host containing pMA1023 and the desired recombinant phage was identified by its ability to confer a Ura^+ phenotype on a lysogen of the *E. coli* K12 *pyrF* mutant strain DB6656 (ROSE, GRISAFI and BOTSTEIN 1984).

Insertion mutagenesis with mini-Tn*10*s: Three yeast genes, *LEU2, RAD50, CDC48,* cloned into pBR322 derivatives, were used as targets for mini-Tn*10* mutagenesis. The target plasmids were pNKY1013 (*RAD50, ARS1, CEN3*) (W. RAYMOND, unpublished data); pRB547 (*CDC48, ARS1:*) (K. FROEHLICH, unpublished data); and YEp21 (*LEU2, ORI-2μ*) (BOTSTEIN *et al.* 1979) for Tn*10*-LUK insertions. pRB38 (*LEU2, URA3, ORI-2μ*) (M. ROSE, unpublished data) was for Tn*10*-LK insertions.

Insertions were isolated according to the standard "λ hop" protocol detailed by WAY *et al.* 1984. Monomeric target plasmids were introduced into *E. coli* strain NK5830 (*recA56 su° lacproXIII*, Arg^-, Ara^-, Nal^R, Rif^R/F′ *laci^q* L8 *pro*) (FOSTER *et al.* 1981) carrying a pACYC184 derivative, pNK629, that produces a high, IPTG-inducible level of Tn*10* transposase function (WAY *et al.* 1984). pNK629 carries a selectable Tet^R marker and is compatible with pBR322-derived plasmids. Pure preparations of monomeric target DNA were obtained by transforming the target plasmid into a $RecA^-$ host such as NK5830 and preparing plasmid DNA from several transformants. The occasional dimeric transformant DNA is readily distinguished from monomeric transformant DNA by its gel mobility and can be discarded. Strains carrying both the target plasmid and the pACYC transposase plasmid were grown to log phase in the presence of 1 mM IPTG, infected with λ mini-Tn*10*-LK or λ mini-Tn*10*-LUK at a multiplicity of 0.4 or less, incubated for 90 min at 37° to allow expression of *kanR* genes, and plated on medium selective for kanamycin resistance. Kan^R colonies result almost exclusively from transpositions of the mini-Tn*10* element from the phage genome into a stable replicon, either the bacterial chromosome, the F′ element or one of the multicopy plasmids; in any host lacking a nonsense suppressor, lysogenization and lytic growth of the phage are prevented by the cI^- Int^- and P^- phenotypes of the phage.

Transpositions into the pBR322 target replicon were specifically obtained from the total collection of transposition events by pooling Kan^R colonies, preparing plasmid DNA by either CsCl or mini-prep methods, transforming a second bacterial strain, and selecting for Kan^R and Amp^R. The secondary host must carry a mutation that prevents adsorption and growth of any λ phage genomes present in the DNA preparations; a similar effect can probably be achieved by incorporation of 1.25 mM sodium pyrophosphate into the plating medium.

Insertion mutations in the desired yeast target gene or region were then obtained by three different approaches:

1. The yeast *LEU2* gene is able to complement *E. coli leuB* mutations (RATZKIN and CARBON 1977) so insertions into *LEU2* were identified by pooling Kan^R transformants, preparing DNA from the pools, introducing that DNA into a $LeuB^-$ bacterial strain (DB1328; λ^R, *leuB proA2, recA*) and screening Kan^R or Amp^R transformants for growth on minimal medium lacking leucine. Plasmids from individual Leu^- colonies were purified for further analysis and transformation into yeast.

Cloned *RAD50* and *CDC48* genes have no scorable phenotype in *E. coli*, so insertions in these genes were identified by introduction of mutagenized DNA back into yeast.

2. In the case of *RAD50*, individual Amp^R Kan^R Tet^S transformants in the secondary host each contained a single insertion in the *RAD50* plasmid and lacked the transposase plasmid. (Fewer than half of the Amp^R and Kan^R transformants from each pool were Tet^R.) Such transformants were analyzed by restriction analysis; plasmids containing insertions anywhere in the cloned yeast segment were introduced into yeast and scored for their ability to complement a haploid *rad50-1* strain, NKY173, for resistance to MMS (0.017% in YPD medium). NKY173 = RM53-3D from R. E. ESPOSITO, genotype *a ade2-1 ade5 can1 his7-2 leu1-12 lys2-1 rad50-1 tyr1-1 ura3-3,13*.

3. In the case of *CDC48*, plasmid DNA was prepared from large pools of Kan^R transformants and transformed into the yeast strain DBY2030 (*cdc48-1 ura3-52*). Transformants were subsequently screened for the *cdc48* cold-sensitive phenotype.

DNA sequence analysis of RAD50::mini-Tn*10*-LUK insertions: The cloned *RAD50* segment on pNKY1013 has a *Sal*I site at each end, and the terminal Tn*10* sequences at each end of the mini-transposon contain an *Nde*I site. For each *RAD50*::mini-Tn*10* insertion, one of the *Nde*I-*Sal*I fragments containing a transposon/target DNA junction was cloned into *Nde*I/*Sal*I digested pNK1252, a derivative of pGC2 (MEYERS and MANIATIS 1985) containing an *Nde*I linker inserted at the *Eco*RI site of the polylinker segment. The resulting plasmids permit direct preparation of single-stranded DNA and easy dideoxy sequence analysis of the cloned segment beginning at a standard oligonucleotide primer site located just beyond the polylinker and extending across the 60 basepairs (bp) of Tn*10* sequence and into adjacent *RAD50* sequences. For most of the insertions analyzed, only one *RAD50*::Tn*10* junction was sequenced; in the three cases where both junctions were sequenced, the 9-bp repeat of target DNA sequences expected for Tn*10* insertions (KLECKNER 1979) was observed. Details of DNA sequence analysis are described by SANGER *et al.* (1980) and MESSING (1983).

Western blot analysis of *LEU2*-β-galactosidase fusion proteins: *E. coli* strain DB1332 (*lacZΔM15*) was transformed with plasmids containing *LEU2* with mini-Tn*10* insertions (see RESULTS). Cultures were grown in LB plus antibiotic until late log phase and the cells from 1.5 ml collected by centrifugation. The cells were immediately resuspended in 75 μl of SDS-sample buffer (LAEMMLI 1970), and boiled for 5 min. Cultures of yeast strain DBY2055 (*ura3-52*) carrying the same plasmids were grown in selective medium (SD minus uracil) to late log-phase. An aliquot of 10 ml of culture was pelleted, and the drained cell pellet was vortexed for 1 min at top speed with an equal volume of glass beads. SDS-sample buffer minus glycerol and running dye was added, and the samples were boiled for 5 min. Beads and debris were removed by centrifugation and glycerol and running dye were added to the supernatant. The bacterial and yeast samples were run on a 6% SDS polyacrylamide gel (LAEMMLI 1970), transferred to nitrocellulose and probed with anti-β-galactosidase (from Cappel) followed by ^{125}I-protein A (from Amersham) as described by BUR-

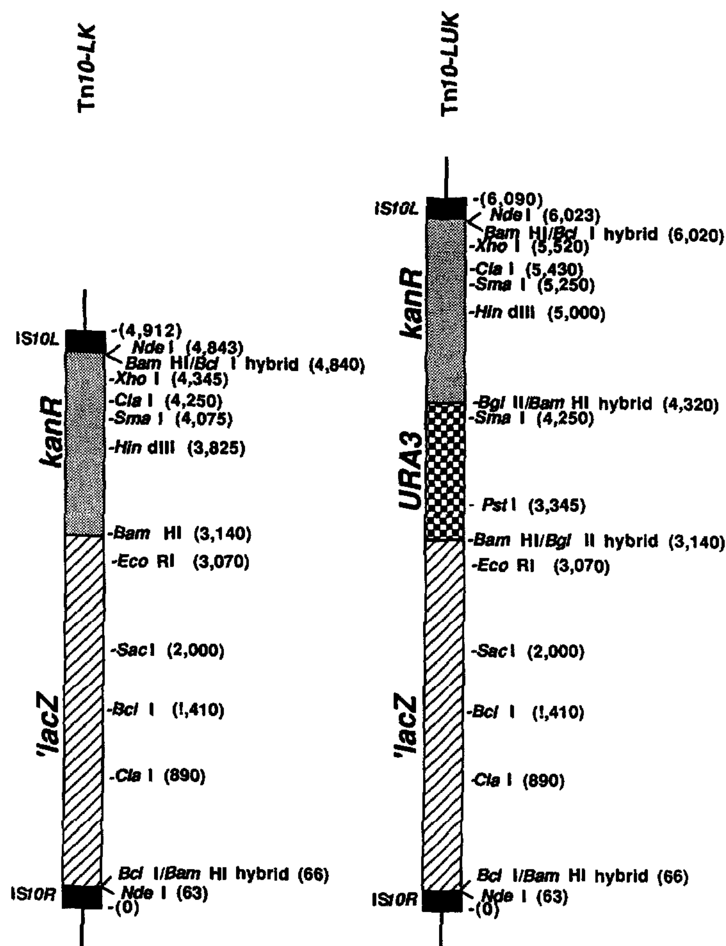

FIGURE 1. Structures of Tn*10*-LK and Tn*10*-LUK. Numbers represent the distance in nucleotides of each restriction enzyme cut site from the left-most end of the element. The orientations of both *lacZ* transcription and *URA3* are from left to right. The sequence of the Tn*10* across the junction to the truncated *lacZ* gene, beginning with bpl of the IS10 sequence, is CTG ATG AAT CCC CTA ATG ATT TTG GTA AAA ATC ATT AAG TTA AGG TGG ATA CAC ATC TTG TCA TAT GAT CCC GTC, etc. The underlined GAT is the first codon of the *lacZ* coding sequence. See text for details of construction.

NETTE (1981). The anti-β-galactosidase was used at a 1:800 dilution in the presence of a glass bead extract of the host yeast cells to reduce nonspecific antibody binding.

RESULTS AND DISCUSSION

Structure of mini-Tn*10*-LK and mini-Tn*10*-LUK fusion hopper transposons: The structure of mini-Tn*10*-LUK and its progenitor, mini-Tn*10*-LK, are shown in Figure 1. Previous work has shown that only the outermost 70 bp of Tn*10* are essential for transposition of the element if Tn*10* transposase function is provided in *trans* (FOSTER *et al.* 1981). We have taken advantage of this situation to replace the central part of the transposon with useful foreign genetic determinants: the *lacZ* structural gene from *E. coli*, the *kanR* determinant from Tn*903*, and in mini-Tn*10*-LUK, the *URA3* gene from *S. cerevisiae*. The *kanR* determinant is present to allow selection for transposition of this element in *E. coli*. The *URA3* determinant allows for direct selection of plasmids or DNA segments containing the transposon in *ura3* yeast strains and in *pyrF E. coli* strains. Mini-Tn*10* derivatives containing other yeast markers can be made in a single step by inserting the desired determinant at the unique *Bam*HI site in Tn*10*-LK. The *lacZ* structural gene segment begins at codon eight of the protein coding sequence and hence lacks both transcriptional and translational initiation signals. An open reading frame extends across one 70-bp terminus of the mini-transposon and into the *lacZ* gene (Figure 1). Insertion of this element into an actively expressed protein coding sequence in the appropriate orientation and reading frame should, in the simplest case, generate an active yeast/β-galactosidase protein fusion.

To facilitate isolation of transposon insertions, the marked mini-transposons are available on bacteriophage λ derivatives suitable for mutagenesis of the bacterial chromosome, plasmids, or episomes as described in detail below; in addition, the elements are available on a pBR322-based multicopy plasmid suitable for mutagenesis of genes present on bacteriophage λ vectors by methods analogous to those described by SNYDER, ELLEDGE and DAVIS (1986). All vectors are described in detail in MATERIALS AND METHODS.

Approaches to isolation and identification of transposon insertion mutations on plasmids. Insertions of a mini-Tn*10*-LUK element into a yeast gene present on a multicopy *E. coli* plasmid can be isolated and identified in two steps. In the first step, large and/

or many independent collections of transposon insertions into the desired target plasmid are generated; in the second step, those pools are screened in *E. coli* and/or in yeast to identify insertions in the gene of interest.

Large numbers of transposon insertions into the target plasmid are isolated as follows. First, a λ phage vehicle carrying the transposon is infected into an su^{o} *E. coli* strain harboring the desired pBR322-derived target plasmid and a compatible pACYC184 plasmid that produces Tn*10* transposase protein. When infected cells are plated on kanamycin selective medium, Kan^R colonies arise at the frequency of about 1 per 10^4 infecting phage genomes. Each such Kan^R colony contains a mini-transposon insertion into the bacterial chromosome, the F′, or one of the multicopy plasmids. A pool of insertions into the target plasmid, mixed with non-transposon containing plasmid molecules, is obtained by scraping 10^3 or more Kan^R colonies off of the selective plates, resuspending them in broth, and extracting plasmid DNA (with or without a short period of growth). In such a DNA preparation, approximately 2% of the target plasmid molecules contain a transposon insertion, as judged by transformation and selection for Amp^R followed by screening for kanamycin resistance; in our analysis of *LEU2*, *RAD50* and *CDC48* genes, insertions into target plasmids of 12.6, 10.7, and 11.1 kilobases (kb) were obtained at frequencies of 2%, 2% and 3%, respectively. From these results we infer that probably about 2% of the original transposon insertions were into the target plasmid and 98% were into the host chromosome, the F′, or the transposase plasmid.

Transposon insertions into the gene or region of interest can be identified by screening of the mixed plasmid DNA pools in one of several ways.

When the target gene has a phenotype in *E. coli*, as is the case for *LEU2*, DNA can be transformed into *E. coli*, target plasmids containing the element identified by selection for Kan^R and Amp^R, and the resulting transformants screened for the relevant phenotypes. In two separate experiments with two different *LEU2* target plasmids (see MATERIALS AND METHODS), about 2% of Kan^R Amp^R transformants from large pools (5 of 268 and 8 of 618) were Leu^-.

If no screen for the yeast gene phenotype is available in *E. coli*, inserts into yeast DNA can be identified in any of three ways:

1. If inserts both inside and outside the gene of interest are desired, the plasmids present in the Kan^R Amp^R (Tet^S) transformants from the DNA pools can be individually characterized by rapid plasmid preparation and restriction enzyme analysis to eliminate insertions into the plasmid backbone. The proportion of insertions into the target region is usually high enough to make such analysis feasible. This approach was used to define the limits of the *RAD50* gene on a 4-kb segment. DNA was extracted from 56 independent pools, each containing about 1000 Kan^R Amp^R transformants, and a total of 49 insertions into the cloned *RAD50* segment was obtained out of the 154 colonies analyzed. The remaining insertions were in the 7.7-kb plasmid backbone.

2. If the target plasmid can be propagated in yeast, insertions into the gene of interest can be identified directly in yeast by transforming plasmid DNA pools into a suitable *ura3*$^-$ yeast strain and screening Ura^+ transformants for the relevant phenotype. Due to the relative inefficiency of yeast transformation, it is recommended that primary DNA pools be used to obtain Kan^R Amp^R transformants in *E. coli* from which DNA is again extracted and used to transform yeast. This approach was used for analysis of *CDC48*. Plasmid DNA from pooled Kan^R Amp^R transformants was extracted and transformed into the appropriate yeast strain. Among 100 Ura^+ yeast transformants, 71 of 100 contained an insertion into the *CDC48* gene. This unusually high frequency is probably a consequence of the fact that the high copy number plasmid carrying the cloned wild type *CDC48* gene is nearly lethal in yeast, so that plasmids containing insertions disrupting *CDC48* function are preferentially recovered.

3. A final method, not employed here, is the selection of transposon insertions into a cloned yeast DNA segment in yeast by gene "transplacement." This method takes advantage of the fact that yeast is very efficiently transformed by linear pieces of chromosomal DNA (ROTHSTEIN 1983). A suitable yeast strain must be transformed with amplified DNA from Kan^R Amp^R colonies that has been cleaved by restriction enzymes within the yeast sequences flanking the region where transposon insertions are desired. This will generate linear fragments containing the target yeast gene, with or without a transposon insertion. In this case, selection for *URA3* will result in substitution of the wild-type chromosomal information with the corresponding information from a linear segment that does carry a transposon insertion.

Locations and phenotypes of mini-Tn*10* insertions in yeast target segments: The physical locations, orientations, and genetic phenotypes of mini-Tn*10* insertions into each of the target genes is shown in Figure 2. The most important conclusion from these data is that Tn*10* insertion into yeast DNA is sufficiently random that the functional limits of even a large gene can easily be defined to the resolution of ±100–200 bp after isolation of 50–75 insertions. In all three cases examined, insertions were obtained at a large number of different sites. Specifically, 49 *RAD50* insertions defined 15 different insertion sites whose distinctness and precise locations in the gene are known from DNA sequence analysis (see below).

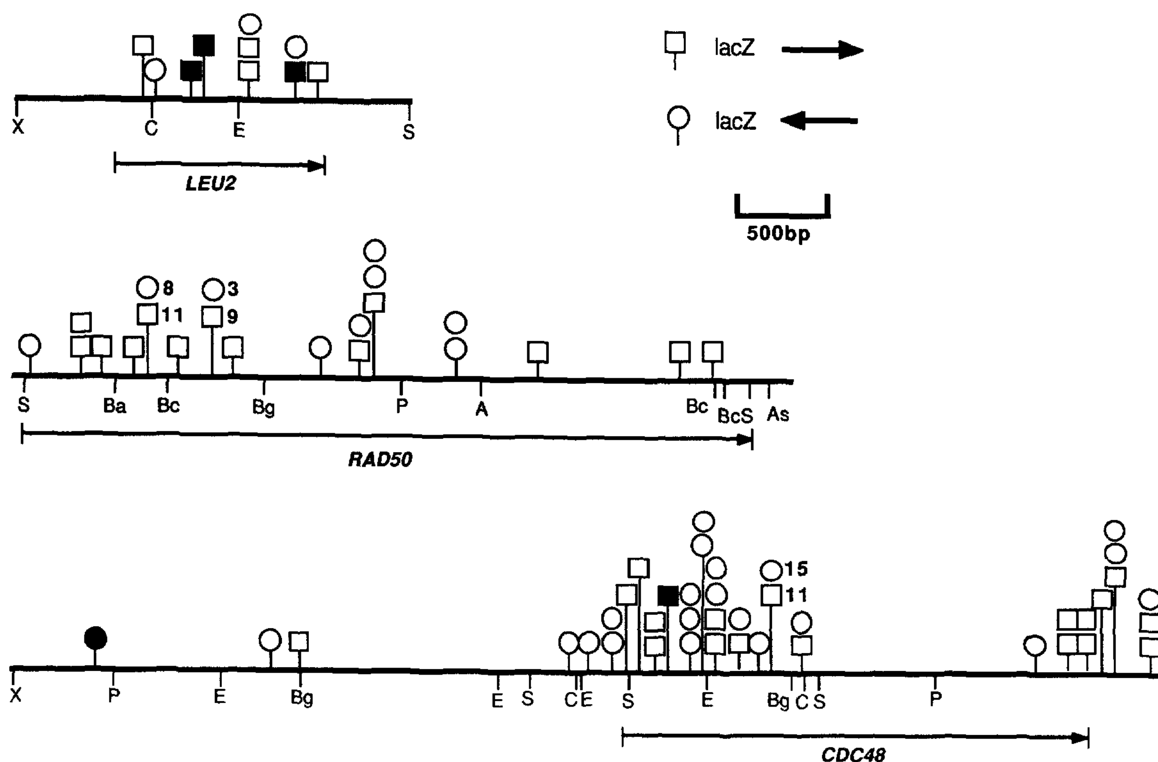

FIGURE 2.—Sites of mini-Tn*10* insertions into yeast DNA. Vertical lines, capped with a circle or a square symbol, represent the mapped sites of Tn*10*-LK and Tn*10*-LUK insertions into the yeast *LEU2*, *RAD50* and *CDC48* genes. The boundaries of *LEU2* and *RAD50* are known from their DNA sequences. The *CDC48* boundaries are defined by the insertions that destroy complementing activity. The two orientations of the inserted elements are depicted by circles and squares, respectively. Numbers next to the symbols represent multiple insertions of the same type. Filled in circles and squares represent insertions that yield a dark blue colony color when transformed into yeast on a multicopy plasmid. Restriction enzyme code: E, *Eco*RI; X, *Xho*I; C, *Cla*I; S, *Sal*I; P, *Pvu*II; Bg, *Bgl*II; H, *Hae*II; Ba, *Bal*I; Bc, *Bcl*I; A, *Ava*II; As, *Asu*II.

Fifty *CDC48* insertions defined a minimum of 14 sites within the gene distinguishable by restriction mapping; other insertions outside of the gene were also well distributed over the remaining 7300 bp of target DNA. Finally, the smaller number (10) of insertions in the *LEU2* gene occurred at a minimum of seven different locations. Since Tn*10* does insert preferentially at particular "hot spots" (see below), there will be occasional target regions where insertion is dominated by a single very hot site. In these situations, Tn*3*-based transposon vectors (SEIFERT *et al.* 1986) which have a less restricted insertion specificity, may be more appropriate.

Previous analysis of Tn*10* insertion into bacterial DNA has shown that the element preferentially inserts into particular "hot spots" and that these sites are all relatives of a single common "consensus" sequence (HALLING and KLECKNER 1982). DNA sequence analysis of mini-Tn*10*-LUK insertions into *RAD50* demonstrates that Tn*10* insertion into yeast DNA follows the same rules as insertion into bacterial DNA. Tn*10* insertion generates a duplication of a 9-bp target site sequence, and the insertion specificity consensus sequence is a symmetrical 6-bp sequence located within the 9-bp repeat sequence (HALLING and KLECKNER 1982). This symmetrical 6-bp sequence is composed of two symmetrically related 3-bp half sites. Figure 3A lists the positions and 9-bp repeat sequences for insertions at each of the 15 identified sites in *RAD50*. Figure 3B summarizes separately the bases found at each position in each of the 30 corresponding *RAD50* specificity sequence half sites and at each of the 22 half sites of the 11 known Tn*10* insertion sites into bacterial DNA. The representation of different basepairs at each position in the half site is very similar for DNA from the two sources. The similarities extend to an additional level as well. For both bacterial and yeast DNA, insertion sites at defined hot spots, where more than a single insertion has been isolated in the standard experimental sample of 20–100 insertions, correspond closely to the consensus sequence; insertions at "not-so-hot" spots, where only a single insertion has been isolated, correspond less closely (Figure 3B).

The rules for Tn*10* insertion specificity suggest that DNA with a higher A+T composition, such as that of *S. cerevisiae*, should contain a higher ratio of "not-so-hot" to "hot" spots than bacterial DNA. The perfect insertion specificity consensus sequence contains four GC and two AT basepairs. However, the allowed deviations from consensus observed at positions 1 and 2 of the "not-so-hot" spots are changes from consensus

A

SITE	LOCATION	NUMBER OF INSERTIONS ORI I	ORI II	TOT	SEQUENCE OF 9-BASEPAIR REPEAT
1	88	0	1	1	G Ⓖ Ⓒ A T G A G C
2	381	2	0	2	T A T Ⓣ C Ⓐ Ⓖ T T
3	498	1	0	1	C Ⓖ Ⓒ A C Ⓐ A G T
4	693	1	0	1	T A T C A Ⓐ Ⓖ T T
5	758	11	8	19	A Ⓖ Ⓒ Ⓣ C Ⓐ A T A
6	895	1	0	1	T C Ⓒ A A Ⓐ Ⓖ T A
7	1100	9	3	12	T Ⓖ Ⓒ A G Ⓐ Ⓖ Ⓒ C
8	1208	1	0	1	A A Ⓒ A T Ⓐ A G T
9	1683	0	1	1	C Ⓖ Ⓒ C A Ⓐ A Ⓒ T
10	1911	1	1	2	C Ⓖ Ⓒ Ⓣ T T Ⓖ T A
11	1985	1	2	3	T A Ⓒ Ⓣ A Ⓐ Ⓖ Ⓒ G
12	2451	0	2	2	A Ⓖ Ⓒ Ⓣ G Ⓐ Ⓖ T C
13	2879	1	0	1	C Ⓖ Ⓒ A A Ⓐ Ⓖ T G
14	3668	1	0	1	A Ⓖ Ⓒ A G G A T G
15	3887	1	0	1	A Ⓖ Ⓒ Ⓣ A Ⓐ T Ⓒ G
		INSERTION SPECIFICITY CONSENSUS SEQUENCE:			5′-N G C T N A G C N N C G A N T C G N-5′

B

CONSENSUS 1 2 3	HALF-SITE POSITION BASE	*E. coli* TOTAL 1	2	3	YEAST TOTAL 1	2	3	*E. coli* HOT SPOT 1	2	3	*E. coli* NON-HOT SPOT 1	2	3	YEAST HOT SPOT 1	2	3	YEAST NON-HOT SPOT 1	2	3
G – –	G	14	0	1	14	0	0	9	0	0	5	0	1	6	0	0	8	0	0
– – –	A	7	0	4	12	1	8	2	0	0	5	0	4	6	0	2	6	1	6
– – T	T	0	3	15	0	8	18	0	1	12	0	2	3	0	2	10	0	6	8
– C –	C	1	19	2	4	21	4	1	11	0	0	8	2	0	10	0	4	11	4

FIGURE 3. Specificity of mini-Tn*10* insertion into yeast *RAD50* DNA. (A) Numbers, locations and 9-bp repeat sequences of 49 *RAD50*::Tn*10*LUK insertions distributed among 15 different sites. Location designated is position of the first basepair of the 9-bp repeat within the sequenced *RAD50* region (E. ALANI and N. KLECKNER, unpublished data). In OriI, the *lacZ* gene and the *RAD50* gene have the same orientation; in OriII, the two genes are oppositely oriented. Sequences are written 5′ to 3′. Circled bases are those where the target site sequence corresponds to the Tn*10* insertion specificity consensus sequence (shown at bottom). (B) Occurrence of different bases at each position in the half sites of identified Tn*10* insertion site specificity sequences. The set of 22 *E. coli* half sites are from the 11 insertion sites analyzed by HALLING and KLECKNER (1982). The set of 30 yeast half sites are from the 15 insertion sites in *RAD50* (Figure 3A). "Hot spots" are sites at which more than a single insertion was identified in an experimental sample of 30–100 insertions; "non-hot-spots" are sites at which only a single insertion was identified.

GC or CG basepairs to nonconsensus AT or TA basepairs (Figure 3B). These considerations suggest that Tn*10* insertions should be more randomly distributed in the typical yeast gene than in the typical bacterial gene. Available data from our most heavily analyzed regions in each organism are consistent with this possibility. One hundred insertions into the 10-kb Salmonella histidine operon are distributed among 19 distinguishable sites (KLECKNER *et al.* 1979): half that number of insertions (49) into a yeast gene half the size (the 4-kb *RAD50* gene) are distributed among nearly the same number (15) of sites (this work).

Tn*10* insertion into yeast DNA exhibits one additional feature characteristic of insertion into bacterial DNA: the element inserts equally frequently in both orientations relative to the target DNA sequence, even for insertions at a single hot spot (see Figure 2).

Use of mini-Tn*10*-LUK insertions to create yeast/lacZ protein fusions: We anticipated that insertions of the mini-Tn*10*-LUK transposon would provide protein fusions between the affected yeast gene and *lacZ*, thus providing a useful tool for analysis of gene expression and protein localization. Insertion in the proper orientation and reading frame within the target gene will generate a fusion having no translational stop signals between the target gene sequences and the truncated *lacZ* gene. This reading frame does contain two ATG translational start codons within the terminal Tn*10* sequence that could potentially have resulted in unwanted expression of β-galactosidase. However, our results show that little detectable *lacZ* expression occurs from these internal start codons in yeast, as expected from the strong preference of yeast translational machinery for the first ATG codon in an mRNA molecule (SHERMAN and STEWART 1982). The situation in *E. coli*, however, is more complex (see below).

In-frame mini-Tn*10*-LK insertions into *LEU2* generate the expected *LEU2-lacZ* protein fusions. Of the seven *LEU2*::Tn*10*-LK insertions having the correct relative orientation of *LEU2* and *lacZ*, three resulted in efficient production of β-galactosidase in both *E. coli* and yeast as judged by intense blue colony color on XGal indicator plates. Western blot analysis, using antibodies to β-galactosidase, of two such insertions (numbers 2 and 4, Figure 4) revealed the presence in both organisms of a β-galactosidase-containing polypeptide having a higher molecular weight than native β-galactosidase. In both cases, the increase in size corresponds to that predicted for a *LEU2-lacZ* fusion from the physical position of the mini-Tn*10* insertion. The other three insertions analyzed (numbers 1, 3 and 5) produced neither dark blue colony color nor reactive polypeptide in yeast, suggesting that the internal ATG codons are not efficiently recognized and that only in-frame insertions result in high levels of β-galactosidase expression in yeast. For insertions 3 and 5, a faint blue color developed in dense yeast patches

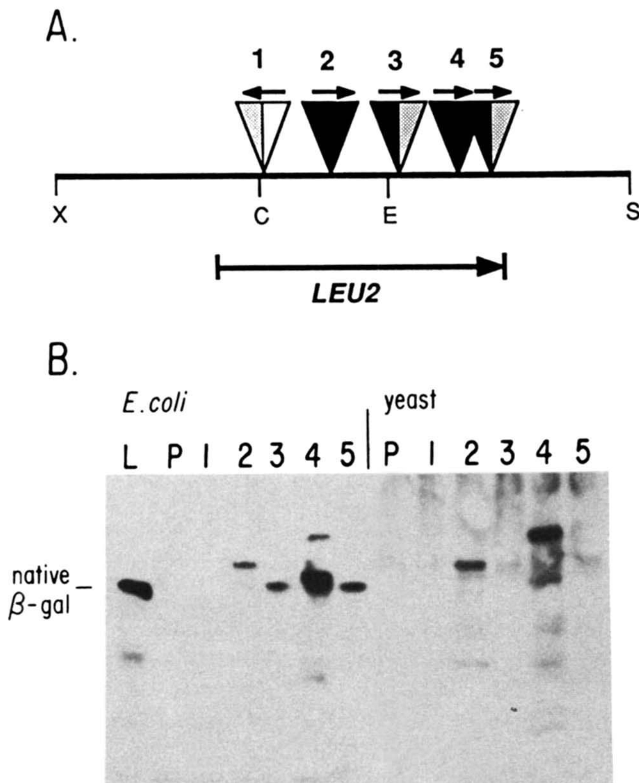

FIGURE 4.—Anti-β-galactosidase "Western blot" analysis of fusion proteins generated by mini-Tn*10* insertions into *LEU2*. (A) Sites and phenotypes of the mini-Tn*10* insertions in *LEU2* analyzed. The arrow above each insertion site indicates the orientation of the element with respect to *lacZ* transcription. The left side of each triangle represents the colony color on X-gal plates of *E. coli* strains carrying these plasmids; the right side represents the color conferred upon yeast. White, no blue color detectable; stippled, faint blue color; filled, dark blue color. (See Figure 2 for restriction enzyme code.) (B) Protein extracts from relevant *E. coli* and yeast strains were run on a polyacrylamide gel, transferred to nitrocellulose and probed with anti-β-galactosidase. Lane L, a Lac^+ strain (no plasmid) induced for β-galactosidase. Lanes P, hosts carrying only the parent plasmid pRB38. Lanes 1 through 5, hosts carrying mini-Tn*10* insertions numbers 1 through 5 on pRB38.

only after prolonged incubation, and little reactive material is evident on the Western blot. In *E. coli*, however, these out-of-frame insertions in the correct orientation produced significant β-galactosidase activity and a reactive polypeptide having the same molecular weight as native β-galactosidase. The most likely explanation for these observations is that in *E. coli*, where rules for translation initiation are different from yeast, *lacZ* translation is sometimes initiated at the internal ATG start codon(s) within the transposon. No blue color was observed in yeast for any insert in the incorrect transcriptional orientation (*i.e.*, insertion number 1).

DNA sequence analysis of the *RAD50*::Tn*10*-LUK insertions showed that only 1 of the 11 sites containing insertions in the correct orientation was present in the correct reading frame. This insertion did not yield blue colony color when introduced into the chromosomal *RAD50* locus, almost certainly because *RAD50* protein is expressed at very low levels in mitotic cells; there is less than 1 transcript per cell of *RAD50* mRNA in mitotically growing yeast cultures (W. RAYMOND, unpublished results).

In the case of *CDC48*, only 1 of the 24 mini-Tn*10* insertions in the correct orientation gave rise to dark blue colonies when introduced into yeast on a high copy number vector; further analysis is required to determine if this is the only in-frame fusion and/or whether it is in some way aberrant or atypical. All of the remaining correctly oriented insertions turned faint blue with prolonged incubation (3 weeks), while incorrectly oriented insertions did not.

From the above results we conclude that, when gene expression is examined in yeast, the mini-Tn*10* elements do generate proper protein fusions and lead to detectable β-galactosidase activity when the target gene is transcribed at an adequately high level. Also, for a target gene like *LEU2* where a significant proportion of insertions give blue color in yeast, the orientation of the target gene can be assigned with a high degree of certainty as being that in which some of the transposon insertions express β-galactosidase.

Expression of β-galactosidase from plasmids bearing yeast::mini-Tn*10*-LUK insertions in *E. coli* requires further comment. All of the *LEU2*, *RAD50* and *CDC48* insertions giving blue color in yeast also gave dark blue color in *E. coli*. However, almost all of the insertions in each of these genes gave at least a small amount of blue color in an *E. coli* strain deleted for the chromosomal *lacZ* gene. For insertions in the wrong orientation or the wrong reading frame, this expression almost certainly results from nonspecific transcription of the gene followed by translation initiated at one of the ATG start codons in the terminal Tn*10* sequences. Variations in colony color from one insertion to another reflect differences in the amount of such nonspecific transcription. Despite these variations, one further result suggests that colony color in *E. coli* can provide circumstantial evidence as to the orientation of the target coding region: all four of the insertions that gave no detectable blue color in *E. coli*, three in *RAD50* and one in *LEU2*, contained insertions in the orientation opposite that of the target coding region. This correlation can be readily rationalized as a "polarity" effect due to coupling between transcription and translation and the effects of nonsense mutations on translation in bacteria. Because evaluations of blue and white colony color vary from one experimenter to another, reference insertions and a standard *E. coli* host strain will be provided upon request.

Application of mini-Tn*10*-LK and -LUK to transposon mutagenesis in bacteria: The mini-Tn*10* vehicles described here can be used effectively to analyze bacterial genes as well as yeast genes. In this case,

however, a caution applies to *lacZ* fusion analysis. Inframe fusions will generate proper fusion proteins. If the gene in question is translated efficiently, the fusion protein should represent the majority of the β-galactosidase activity expressed, and the fusion should be appropriate for all standard applications. A small amount of translation from ATG start codons within the Tn*10* terminal sequences should be anticipated but should not interfere.

In two situations, however, the internal start codons can be expected to cause confusion. First, insertions in the wrong reading frame may result in expression of β-galactosidase that is properly regulated at the transcriptional level but does not result from translation initiation at the target gene initiation signals. Second, if the target gene is translated very inefficiently, translation from the internal start codons may be sufficient to mask normal translation even from inframe fusions. In both of these situations, the insertions obtained retain their usefulness as "operon" fusions rather than "protein" fusions and can be used in any case to study transcriptional regulation in the usual way.

Other advantages of mini-Tn*10* insertion analysis: Once isolated and characterized, a collection of mini-transposon insertions in a gene or region of interest provides a valuable set of tools for further genetic and physical analysis. The presence of the *URA3* marker on the Tn*10*-LUK element means that insertions can be used directly to construct chromosomal gene disruptions in yeast. The insertions can also be used in standard ways to provide correlations between genetic and physical maps of the target region. It is particularly worth noting that the insertions provide convenient starting points for DNA sequence analysis of the target region. Most simply, junction fragments extending from an *Nde*I site in the terminal Tn*10* sequences (Figure 1) to a suitable site outside of the relevant segment can be reiteratively subcloned into a hybrid single-stranded phage/plasmid vector. DNA sequence analysis can then be carried out either by having juxtaposed the NdeI site to a standard primer binding site in the vector or by using an oligonucleotide primer homologous to the terminal Tn*10* sequences. The former approach was used in sequence analysis of the *RAD50* insertions (Figure 3; MATERIALS AND METHODS).

O.H. was supported by an European Molecular Biology Organization fellowship. M.A.H. was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation. Work by O.H. and P.E. was supported by a grant to N.K. from the National Science Foundation. Work by M.A.H. was supported by grants to D.B. from the American Cancer Society (UC90) and from the National Institutes of Health (GM18973 and GM21253).

LITERATURE CITED

BOTSTEIN, D., S. C. FALCO, S. E. STEWART, M. BRENNAN, S. SCHERER, D. T. STINCHCOMB, K. STRUHL and R. W. DAVIS, 1979 Sterile host yeasts (SHY): a eukaryotic system of biological containment for recombinant DNA experiments. Gene **8:** 17–24.

BURNETTE, W. N., 1981 'Western blotting'; electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112:** 195–203.

CLIFTON, D., S. B. WEINSTOCK and D. G. FRAENKEL, 1978 Glycolysis mutants in *Saccharomyces cerevisiae.* Genetics **88:** 1–11.

FOSTER, T. J., M. A. DAVIS, D. E. ROBERTS, K. TAKESHITA and N. KLECKNER, 1981 Genetic organization of transposon Tn*10*. Cell **23:** 201–213.

HALLING, S. M. and N. KLECKNER, 1982 A symmetrical six-base-pair target site sequence determines Tn*10* specificity. Cell **28:** 155–163.

HALLING, S. M., R. W. SIMONS, J. C. WAY, R. B. WALSH and N. KLECKNER, 1982 DNA sequence organization of Tn*10*'s IS*10*-Right and comparison with IS*10*-Left. Proc. Natl. Acad. Sci. USA **79:** 2608–2612.

HUISMAN, O. and N. KLECKNER, 1987 A new generalizable test for detection of mutations affecting Tn*10* transposition. Genetics **116:** 185–189.

KALNINS, A., K. OTTO, U. REUTHER and B. MULLER-HILL, 1983 Sequence of the *lacZ* gene of *E. coli.* EMBO J. **2:** 593–597.

KLECKNER, N., 1979 DNA sequence analysis of Tn*10* insertions: origin and role of 9-bp flanking repetitions during Tn*10* translocation. Cell **16:** 711–720.

KLECKNER, N., D. STEELE, K. REICHARDT and D. BOTSTEIN, 1979 Specificity of insertion by the translocatable tetracycline-resistance element Tn*10*. Genetics **92:** 1023–1040.

LAEMMLI, U. K., 1970 Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 580–585.

MESSING, J., 1983 New M13 vectors for cloning. Methods Enzymol. **101:** 22–78.

MEYERS, R. M. and T. MANIATIS, 1985 A new method for saturated mutagenesis. Science **229:** 242–245.

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

OKA, A., H. SUGISAKI and M. TAKANAMI, 1981 Nucleotide sequence of the kanamycin resistance transposon Tn903. J. Mol. Biol. **147:** 217–226.

RATZKIN, B. and J. CARBON, 1977 Functional expression of cloned yeast DNA in *Escherichia coli.* Proc. Natl. Acad. Sci. USA **74:** 487–491.

ROSE, M., P. GRISAFI and D. BOTSTEIN, 1984 Structure and function of the yeast *URA3* gene: expression in *Escherichia coli.* Gene **29:** 113–124.

ROTHSTEIN, R. J., 1983 One-step gene disruption of yeast. Methods Enzymol. **101:** 202–211.

SANGER, F., A. R. COULSON, B. G. BARRELL, A. J. H. SMITH and B. A. ROE, 1980 Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:** 161–178.

SEIFERT, H. S., E. Y. CHEN, M. SO and F. HEFFRON, 1986 Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA **83:** 735–739.

SHERMAN, F. and J. W. STEWART, 1982 Mutations altering initiation of translation of yeast iso-1-cytochrome *c*; contrasts between the eukaryotic and prokaryotic initiation process. pp. 301–333. In: *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*, Edited by J. N. Strathern, E. W. Jones and J. R. Broach, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

SHERMAN, F., G. R. FINK and J. B. HICKS, 1983 *Methods in Yeast*

Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

SNYDER, M., S. ELLEDGE and R. W. DAVIS, 1986 Rapid mapping of antigenic coding regions and constructing insertion mutations in yeast genes by mini-Tn*10* 'transplason' mutagenesis. Proc. Natl. Acad. Sci. USA **83:** 730–734.

WAY, J. C., M. A. DAVIS, D. MORISATO, D. E. ROBERTS and N. KLECKNER, 1984 New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32:** 369–379.

Communicating editor: E. W. JONES