

# Artificial mobile DNA element constructed from the *EcoRI* endonuclease gene

SEAN R. EDDY AND LARRY GOLD

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309-0347

Communicated by William B. Wood, November 12, 1991 (received for review October 16, 1991)

**ABSTRACT** There exist several examples of mobile group I introns. These introns appear to use a straightforward mechanism to achieve highly site-specific and efficient insertion into homologous intronless genes. Because the only intron-specific function required by the prevailing model for the mechanism of intron mobility is the introduction of a site-specific double-stranded break in the intronless recipient DNA molecule, we reasoned that it should in principle be possible to construct artificially mobile DNA sequences. We have constructed an artificial mobile element from the gene for the restriction enzyme *EcoRI* that is capable of site-specific insertion at rates near those of authentic mobile introns. The generality of the mobility mechanism may enable high-efficiency targeted gene replacements or disruptions in a variety of organisms.

All of the mobile group I introns encode a protein with double-stranded DNA (dsDNA) endonuclease activity specific for a site near the point of intron insertion in an intronless gene (1). Acquisition of a mobile group I intron shows genetic features that are consistent with a nonreciprocal replicative homologous recombination event, including a dependence on flanking exon homology in both donor and recipient and a coconversion of flanking markers from the donor into the recipient (1–3). The introns are, therefore, thought to move by a mechanism similar to the double-stranded break repair (DSBR) model proposed by Szostak and colleagues (4) to explain aspects of meiotic homologous recombination (1, 4) (Fig. 1).

It is possible that the intron-encoded endonucleases also possess additional activities to optimize the efficiency of the process. For instance, an intron-encoded endonuclease might act to nucleate the assembly of the correct recombination machinery at the broken ends. There are mobile elements that encode multifunctional proteins with *in vitro* dsDNA endonuclease activity, such as retroviral integrases (5), various transposases (6), or the *Bombyx mori* retrotransposon R2Bm, which encodes an rRNA gene-site-specific DNA endonuclease that is thought to mobilize the retrotransposon with a break-join retroviral integrase mechanism (7).

Are the proteins encoded by mobile group I introns simply site-specific DNA endonucleases? Or, do they possess additional activities that increase the efficiency of DSBR event?

As one approach to answering this question, we have constructed an artificial mobile element that uses a “generic” restriction enzyme as its homing endonuclease and asked if the frequency of its mobility approaches that catalyzed by authentic intron-encoded homing endonucleases. We have chosen to use the well-known restriction endonuclease *EcoRI* for these experiments. We have configured the experimental system to be closely analogous to the plasmid-to-phage mobility experiments used to measure mobility frequencies for the mobile phage T4 *td* and *sun Y* introns.

## MATERIALS AND METHODS

**Media and Chemicals.** Cells were grown in  $\lambda$  broth (10 g of Bacto tryptone and 8 g of NaCl per liter of water) supplemented with 0.2% maltose and chloramphenicol (33  $\mu\text{g}/\text{ml}$ ) or ampicillin (50  $\mu\text{g}/\text{ml}$ ) were added as needed to maintain plasmids. Top agar for plating was  $\lambda$  broth plus 0.8% agar. Bottom agar was  $\lambda$  broth plus 1.5% (wt/vol) agar and 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactoside at 40  $\mu\text{g}/\text{ml}$ . SM for phage dilutions is as described by Maniatis *et al.* (8).

**Bacterial and Phage Strains.**  $\lambda\text{gt}11$  DNA for cloning into pL501 was purchased from Stratagene;  $\lambda\text{gt}11$  phage was produced by transfection of this naked DNA into LE392 and subsequently propagated by making plate stocks (9) with LE392 as the host cells. LE392 cells harboring the RI methylase-expressing pSRR101 plasmid were used as hosts to produce methylated  $\lambda\text{gt}11$  phage stocks.

*Escherichia coli* LE392 was provided by W. Studier (Brookhaven National Laboratory, Long Island, NY). AB1157 and its derivatives, JC5519 and JC8679, were obtained from the *E. coli* Genetic Stock Center (CGSC; nos. 1157, 5114, and 6490, respectively). Strains M5160 and M5217 were provided by S. Adhya (National Cancer Institute, Bethesda, MD).

**Plasmids.** The *EcoRI*-bearing plasmid pRI28 was provided to us by G. Wilson of New England Biolabs. pDIP19C is a pACYC184-based vector, compatible with pBR322-based plasmids like pRI-28 in two-plasmid experiments, constructed in this laboratory (10).

Construction of pL501, pSRR101, pLIN101, pLIR201, and pLIR301 was done by standard recombinant DNA techniques (8). The methylase-expressing plasmid pSRR101 was constructed by filling-in and disrupting a unique *Bgl* II site in the coding region for the *EcoRI* endonuclease gene on pRI28. The *EcoRI* site of vector pDIP19C was eliminated by filling-in and a 1.2-kilobase (kb) *Bst*YI-*Kpn* I fragment of  $\lambda\text{gt}11$  spanning the unique *EcoRI* site was subcloned into the pDIP19C *Bam*HI-*Kpn* I sites to create pL501. The 4-base AATT intervening sequence (IVS) in pLIN101 was created by filling-in the *EcoRI* site. A 2.2-kb *Bam*HI-*Cla* I fragment of pRI-28 carrying the *EcoRI* R and M genes under control of their own promoter was subcloned into a blunted *EcoRI* site of pL501 to create the pLIR201; pLIR301 was constructed in the same manner except that the 2.2-kb fragment was lifted from the endonuclease-defective pSRR101 plasmid.

**Plasmid to Phage Mobility.** For *sbca* experiments, plasmid-containing cells were grown to an OD<sub>600</sub> value of 0.6 at 37°C. Cells (7.5 ml) were harvested by low-speed centrifugation, resuspended in 1.3 ml of  $\lambda$  broth plus maltose, and titered (titers are  $\approx 10^9$  cells per ml). In a prewarmed tube at 37°C, 450  $\mu\text{l}$  or 900  $\mu\text{l}$  of this cell suspension was infected with 50  $\mu\text{l}$  or 100  $\mu\text{l}$ , respectively, of  $\lambda\text{gt}11$  diluted to  $10^9$  plaque-forming units/ml. The final multiplicity of the infection is thus 0.1. A sample of a 1:10 dilution of the adsorption tube was killed with

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: dsDNA, double-stranded DNA; DSBR, double-stranded break repair; IVS, intervening sequence.

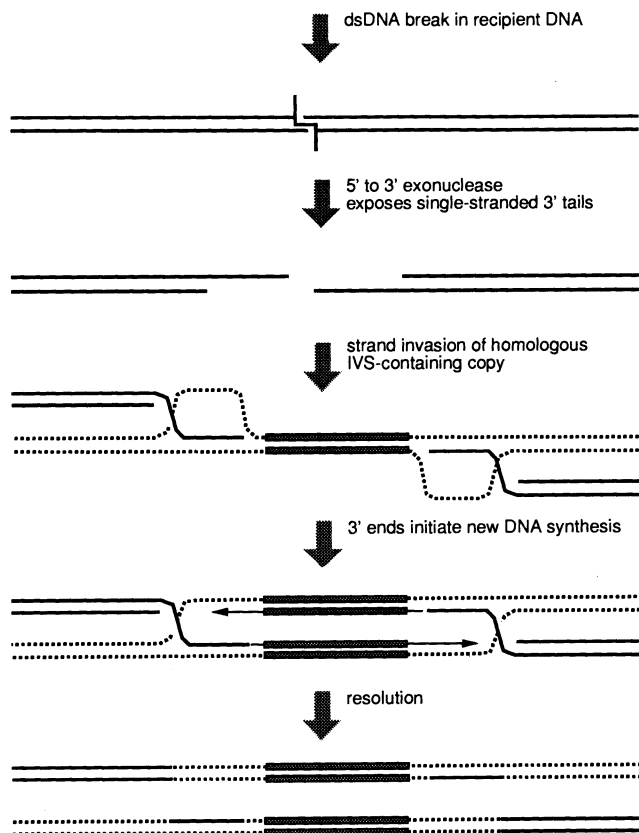


FIG. 1. DSBR model proposed to account for intron mobility (1). The recipient molecule (solid lines) is cut by a specific dsDNA endonuclease, and the break is recombinationally repaired using an unbroken intron-containing copy of the gene (stippled lines) as template.

a drop of chloroform and titered to determine the number of unadsorbed phage 5 min after infection. At 10 min after infection, infected cells were serially diluted 1:1000 to prevent further adsorption. At 90 min after infection, the diluted infected cells were artificially lysed by adding a few drops of chloroform; further dilutions of this sample were titered on  $\lambda/5$ -bromo-4-chloro-3-indolyl  $\beta$ -D-galactoside plates.

*reda $\beta$*  experiments were done identically, except that the cells were grown at 30°C to an OD<sub>600</sub> value of 0.6, titered, shifted up to 39°C for 20 min to induce the promoter of the defective lysogen, and then concentrated by centrifugation; the rest of the infected culture was also then incubated at 39°C rather than 37°C.

In these experiments, actual multiplicities of infection ranged from 0.03 to 0.2, and the percent adsorbed phage ranged from 90 to 99%.

**PCR Assay of Single Plaques.** Single plaques were picked with a sterile Pasteur pipette and resuspended in 100  $\mu$ l of water. A 40- $\mu$ l sample of this was used as template DNA in a 100- $\mu$ l PCR mixture that also contained 50 mM KCl, 10 mM Tris Cl (pH 8.6), 2.5 mM MgCl<sub>2</sub>, bovine serum albumin (170  $\mu$ g/ml), all four dNTPs (each at 400  $\mu$ M), 3 units of *Taq* DNA polymerase (Promega), and the two amplification primers [LAC412, 5'-TGCGGGACGCGGAATTGAA-3', and LAC433, 5'-GACACCAGACCAACTGGTAA-3' (50 pmol each)], which anneal to sequences flanking the *Eco*RI site of  $\lambda$ gt11 (Fig. 2). These were covered with mineral oil and incubated for 50 cycles of 95°C for 1 min, 50°C for 1 min, and 70°C for 4 min, and then 10% of this reaction mixture was electrophoresed on a 1.0% agarose gel, which was stained in ethidium bromide and photographed under short-wave UV illumination.

## RESULTS

The DSBR model, diagrammed in Fig. 1, predicts that a mobile element akin to the mobile group I introns must satisfy the following requirements (1):

(i) A site-specific dsDNA break is introduced in the recipient DNA at or near the insertion site, which in most cases will be directed by an endonuclease encoded by the mobile element. As only a dsDNA break in a homologous gene is utilized, the specificity of this endonuclease may approach one site per genome.

(ii) The donor DNA molecule, containing the mobile IVS, is immune to the action of the homing endonuclease; a simple way to accomplish this is to have the IVS disrupt the recognition site of the endonuclease in the donor copy. If both molecules were to be cut, the unidirectionality of the gene conversion event would be lost.

(iii) The "exon" sequences of the donor and recipient DNA molecules must be homologous to allow for strand invasion.

The recipient DNA in these experiments is the genome of the bacteriophage  $\lambda$ gt11, which contains a single *Eco*RI recognition site in the C terminus of the  $\beta$ -galactosidase gene *lacZ* (Fig. 2) (11). Using a phage genome as the recipient allowed us to mimic the extreme site specificity of natural intron-encoded homing endonucleases; expression of the *Eco*RI methylase from the mobile element (see below) modifies genomic *Eco*RI sites and makes the incoming  $\lambda$ gt11 *Eco*RI site functionally unique.

The mobile *Eco*RI element was constructed by flanking the *Eco*RI restriction and modification genes with "exons" from

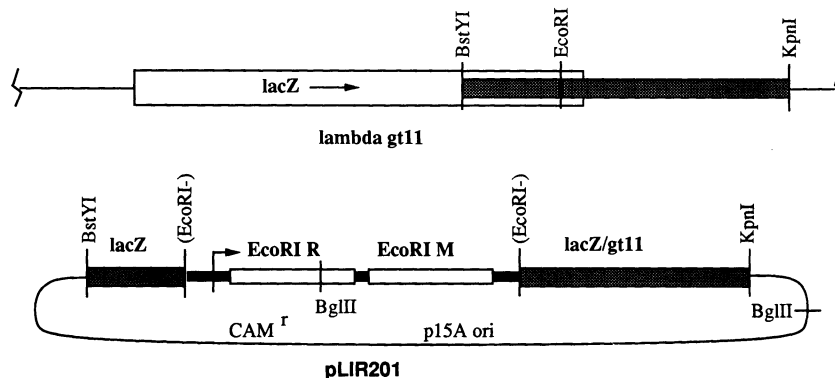


FIG. 2. Recipient and donor DNAs used in this experiment. The shared region of homology between donor and recipient molecules is shaded (260 bp to the left and 915 bp to the right of the *Eco*RI site).  $\lambda$ gt11 tolerates insertions of up to 8.3 kb. The *Eco*RI insertion in pLIR201 is 2.2 kb long, and the entire plasmid is 7.8 kb long.

Table 1. Bacterial strains used in this study

Strain	Relevant genotype	Note(s)	Ref.
LE392	F <sup>-</sup> <i>supE44 supF58 lacY1</i>	Permissive $\lambda$ gt11 host	8
AB1157	F <sup>-</sup>	recBC pathway	16
JC5519	AB1157 <i>recB21 recC22</i>	rec <sup>-</sup>	16
JC8679	AB1157 <i>recB21 recC22 sbcA23</i>	recE pathway	16
M5160	F <sup>-</sup> $\Delta$ ( <i>bio</i> ) [ $\lambda$ cl857 $\Delta$ ( <i>cro-chlA</i> )H1]	Inducible Red pathway	17
M5217	M5160 $\Delta$ ( <i>int-gam</i> )bio11	Red <sup>-</sup> gam <sup>-</sup> ; RecBC pathway	18

the  $\lambda$ gt11 *lacZ* gene. The construction was done so that the *EcoRI* cassette disrupts the *EcoRI* site in *lacZ*. This donor DNA is carried on the plasmid pLIR201 (Fig. 2).

After some failed pilot experiments, we realized from work done by F. W. Stahl and coworkers (12–14) that the type of host recombination system is a key variable. Although the wild-type *E. coli* recBC system is stimulated by the presence of a dsDNA break, this may be because the break is used as an entry point for the translocation activity of recBC (12); the 3' ends at the break are not used as invasive substrates, as is supposed in the model for intron mobility shown in Fig. 1. On the other hand, the  $\lambda$  Red and cryptic *E. coli* RecE pathways do seem to use a dsDNA end as an invasive substrate, possibly because these recombination pathways utilize a 5'–3' exonuclease activity required to efficiently expose long 3' single-stranded DNA overhangs (12–15). We therefore included in our experiments host cells expressing the  $\lambda$  Red pathway from a defective prophage or *recBC*<sup>-</sup> cells carrying the *sbcA* mutation, which activates the RecE pathway (Table 1).

Table 2 shows the results of experiments in which various cells carrying the *EcoRI* construct were infected with  $\lambda$ gt11, and progeny phage were plated. Phage that acquire the *EcoRI* element make clear plaques on 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactoside plates, whereas the parental phage make blue plaques. Ten percent of the progeny from cells expressing the Red recombination pathway and 3% of the progeny from cells expressing the RecE recombination pathway produced clear plaques, indicative of mobility. Eliminating *EcoRI* function by site-directed mutagenesis of the RI gene on the donor plasmid or protecting the  $\lambda$ gt11 *EcoRI* site by methylation abolished mobility, indicating the requirement for the dsDNA break. Elimination of the recombination system or use of the wild-type *E. coli* recBC recombination pathway reduced mobility to background levels. In Fig. 3, several clear progeny plaques were picked and analyzed by the PCR, to ascertain that they did in fact pick up the 2.2-kb insertion.

Earlier pilot experiments in which the donor DNA, rather than containing the 2.2-kb *EcoRI* cassette, instead contained a short 4-base-pair (bp) (AATT) IVS, and *EcoRI* restriction and modification functions were provided in trans from a compatible second plasmid, gave higher frequencies (43% in the Red recombination system and 15% in the RecE system, data not shown). We do not know if this higher frequency reflects the shorter length of the mobile heterology or a higher

expression level of *EcoRI* from the high copy number pBR322 vector used as the second plasmid in the pilot experiments. It was possible in these experiments to exclude the possibility that *EcoRI* restriction forces the phage genome to recombine with the plasmid to survive; average burst sizes were identical within experimental error for infections that either did or did not carry the homologous IVS-containing donor plasmid.

## DISCUSSION

Thaler *et al.* (13) have shown that an *EcoRI*-induced double-strand break in a recipient  $\lambda$  chromosome stimulates the acquisition of a nearby point mutation from homologous plasmid DNA, such that  $\approx$ 9% of the progeny  $\lambda$  acquire the point mutation. Kobayashi and Takahashi (15) have shown that a 283-bp double-stranded gap created by restriction enzyme cutting *in vitro* can be repaired by DSBR from a homologous donor sequence on the same plasmid molecule at high efficiency after transformation of the linearized molecules into *E. coli* cells (15). Our experiments extend these results to a quite large insertion sequence and demonstrate that it is possible to construct a reasonably efficient mobile DNA element that encodes just a site-specific dsDNA endonuclease activity.

We have configured the experimental system to closely mirror that used to measure the mobility of the mobile T4 *td* and *sunY* introns. Bell-Pedersen *et al.* (19) have shown that the *td* intron-encoded homing endonuclease can catalyze the mobility of a foreign gene (for kanamycin resistance or *lacZ*) within the *td* exon context as efficiently as the *td* intron itself. We thus presume that any differences between the mobility of our artificial *EcoRI* element and an authentic mobile group I intron reflect differences between a restriction enzyme and an intron-encoded homing endonuclease.

In plasmid-to- $\lambda$  mobility experiments, the phage T4 *td* intron-encoded endonuclease is seen to move an insertion sequence with a frequency of 14–73% (3, 19). In these experiments, we see that *EcoRI* can move a 2.2-kb insertion with a frequency of 3–10%, which appears to be an  $\approx$ 100-fold enhancement over background acquisition levels. It therefore seems true that a simple endonuclease can function as the homing endonuclease of a mobile DNA element. There remains a 5- or 10-fold difference in frequency to be accounted for, however.

Table 2. Single-cycle  $\lambda$ gt11 infections of cells carrying either the mobile *EcoRI* construct pLIR201 or the RI<sup>-</sup> control pLIR301

Summary	Rec pathway	Plasmid	Host cell	No. clear/no. total plaques	% mobility
Complete system	RecE	pLIR201	JC8679	155/4488	3
No site (me <sup>+</sup> phage)	RecE	pLIR201	JC8679	0/516	<0.2
No <i>EcoRI</i>	RecE	pLIR301	JC8679	1/1221	0.1
recBC <sup>+</sup>	RecBC	pLIR201	AB1157	8/2972	0.3
recBC <sup>-</sup>	None	pLIR201	JC5519	0/771	<0.2
Complete system	Red	pLIR201	M5160	109/1113	10
No site (me <sup>+</sup> phage)	Red	pLIR201	M5160	0/1409	<0.1
No <i>EcoRI</i>	Red	pLIR301	M5160	0/1548	<0.1
Red <sup>-</sup> gam <sup>-</sup>	RecBC	pLIR201	M5217	0/2033	<0.1

Stock  $\lambda$ gt11 had a background level of 10<sup>-3</sup> to 10<sup>-4</sup> clear plaque formers (two clear plaques per 6579 total plaques).

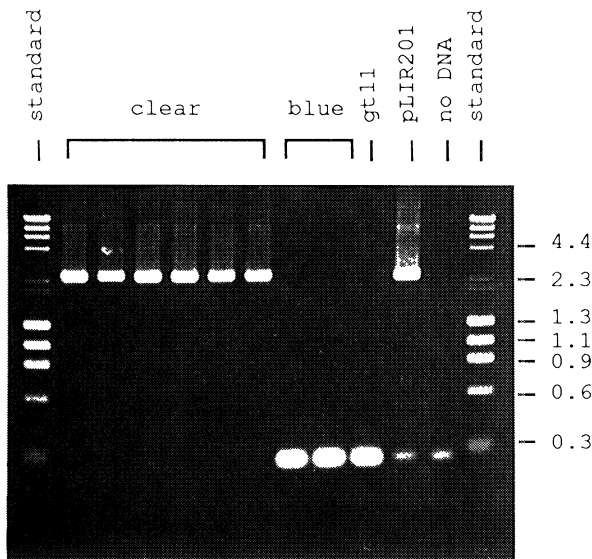


Fig. 3. PCR analysis of single progeny plaque isolates. A 250-bp region including the  $\lambda$ gt11 *Eco*RI site was amplified. Six of six clear plaques showed a 2.4-kb band, indicating the presence of the 2.2-kb mobile *Eco*RI element. Blue plaques from the same plate (two of two) showed the intronless 250-bp fragment, as did the parental  $\lambda$ gt11. This experiment does not distinguish clean insertion of the element, integration of the plasmid (as would occur if DSB intermediate shown in Fig. 1 resolves in a single-crossover configuration), or an apparently "clean" insertion of the element that would result from a two-step event of plasmid integration followed by deletion between the direct repeats created by a plasmid integration event. Sizes of molecular size standards are shown in bp ( $\times 10^{-3}$ ).

One important difference between the behaviors of our artificial *Eco*RI mobile element and the mobile introns is that *Eco*RI digests  $\lambda$ gt11, reducing average burst sizes 10- to 20-fold. Cutting induced by the phage T4 *td* and *sunY* introns does not appear to reduce the burst size of intronless T4 phage (unpublished observations). It has been noted that the broken DNA ends induced by the *Saccharomyces cerevisiae* *al4* intron endonuclease seem to be unusually stable (20). It may be that intron-encoded endonucleases, unlike restriction enzymes, have an additional activity that prevents lethal degradation at the dsDNA break before recombinational repair can take place. This might conceivably contribute to the higher mobility of authentic mobile introns.

Interestingly, the wild-type *E. coli* *recBC* recombination pathway does not support mobility in these experiments. The  $\lambda$  Red or the mutationally activated *E. coli* RecE pathway, both of which are distinguished from the RecBC pathway by the use of 5'-3' exonuclease activity (14), must be established before infection by the recipient chromosome to support the mobility event. This dependence on a particular recombination system may limit the spread of mobile group I introns and may account for some aspects of the distribution of group I introns in the biosphere (in particular, the apparent absence of mobile introns from the *E. coli* genome). It probably also explains the previously puzzling failure of plasmid-to-plasmid mobility experiments (which were carried out in *recBC*<sup>+</sup> cells) (3). We presume that the phage T4 recombination system, which is known to support efficient intron

mobility, must be similar to the Red and RecE pathways rather than the RecBC pathway in this respect.

It should be possible to use DSB-based targeting of homologous recombination to effect site-directed gene conversions in a variety of organisms. The extreme specificity of intron-encoded endonucleases [18 bp in one case (21)] could allow targeting of a single specific site in a genome, either by the good fortune of having a usable site already in the genome or by the one-time transgenic introduction of a "landing site" containing the endonuclease recognition site and flanking exon homology to some shuttle vector construct. A related strategy is to exploit the dsDNA breaks left behind by transposition of certain transposons to effect precise modifications or replacements of the DNA surrounding transposon-containing alleles of a gene; such manipulations have recently been performed using transposable *P* elements in *Drosophila* (22, 23).

We thank C. Tuerk, D. Parma, and S. Ringquist for critical reading of this manuscript; G. Wilson and New England Biolabs for the *Eco*RI clone; and S. Adhya, W. Studier, and B. Bachmann for the bacterial strains. This work was supported by Grants GM19963 and GM28685 from the National Institutes of Health. S.R.E. was supported by a National Science Foundation Graduate Fellowship.

- Dujon, B. (1989) *Gene* **82**, 91-114.
- Bell-Pedersen, D., Quirk, S., Aubrey, M. & Belfort, M. (1989) *Gene* **82**, 119-126.
- Quirk, S., Bell-Pedersen, D. & Belfort, M. (1989) *Cell* **56**, 455-465.
- Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983) *Cell* **33**, 22-35.
- Craigie, R., Fujiwara, T. & Bushman, F. (1990) *Cell* **62**, 829-837.
- Berg, D. E. & Howe, M. M. eds. (1989) *Mobile DNA* (Am. Soc. Microbiol., Washington).
- Xiong, Y. & Eickbush, T. H. (1988) *Cell* **55**, 235-246.
- Maniatis, T., Fritsch, E. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Silhavy, T., Berman, M. & Enquist, L. (1984) *Experiments with Gene Fusions* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Singer, B. & Gold, L. (1991) *Gene* **106**, 1-6.
- Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194-1198.
- Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987) *J. Mol. Biol.* **195**, 75-87.
- Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987) *Genetics* **116**, 501-511.
- Thaler, D. S. & Stahl, F. W. (1988) *Annu. Rev. Genet.* **22**, 169-197.
- Kobayashi, I. & Takahashi, N. (1988) *Genetics* **119**, 751-757.
- Gillen, J. R., Willis, D. K. & Clark, A. J. (1981) *J. Bacteriol.* **145**, 521-532.
- Greer, H. (1975) *Virology* **66**, 589-604.
- Toothman, P. & Herskowitz, I. (1980) *Virology* **102**, 133-146.
- Bell-Pedersen, D., Quirk, S., Clyman, J. & Belfort, M. (1990) *Nucleic Acids Res.* **18**, 3763-3770.
- Delahodde, A., Goguel, V., Becam, A., Creusot, F., Perea, J., Banroques, J. & Jacq, C. (1989) *Cell* **56**, 431-441.
- Colleaux, L., D'Auriol, L., Galibert, F. & Dujon, B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6022-6026.
- Engels, W. R., Johnson-Schlitz, D. M., Eggleston, W. R. & Sved, J. (1990) *Cell* **62**, 515-525.
- Gloor, G. B., Nassif, N. A., Johnson-Schlitz, D. M., Preston, C. R. & Engels, W. R. (1991) *Science* **253**, 1110-1117.