JOURNAL OF BACTERIOLOGY, Apr. 1982, p. 389-394
0021-9193/82/040389-06$02.00/0

Vol. 150, No. 1

DNA Homology of the Promoter-Distal Regions of the *tra* Operons of Sex Factors F and R100 in *Escherichia coli* K-12

PAUL A. MANNING†* AND GIOVANNA MORELLI

Max-Planck-Institut für molekulare Genetik, Abt. Trautner, D-1000 Berlin 33, Federal Republic of Germany

Received 20 April 1981/Accepted 20 November 1981

The promoter-distal regions of the *tra* operons of F and R100-1 were analyzed by heteroduplex analysis, and the regions of nonhomology were identified. A common *Eco*RI restriction site was shown to be present, and this has allowed the physical maps to be aligned.

The sex factor F of *Escherichia coli* K-12 and the drug resistance factor R100 are closely related in a number of their transfer (*tra*) genes (1, 4). It has also been previously shown that there are stretches of the DNA of the two sex factors within the *tra* region which can hybridize, alternating with stretches which cannot (8). Thus, we were interested to compare nonhomologies with the *tra* cistrons on the two physical maps. It was also of interest to see how much of the *tra* operon was carried by the mini-R100 plasmid pSM6, as well as to elucidate how it had been formed from the parental R12 (R100) plasmid. In the two accompanying papers (4, 5), we have been able to construct physical maps of the DNA of the promoter-distal regions of two F-like sex factors, F and R100. In this note we present the data which has enabled us to define the regions of nonhomology between the *tra* regions of F and R100 and to align them with the physical maps.

We used the chimeric plasmids pRS26, pRS29, and pRS31 as our source of F factor DNA (2, 9; Fig. 1). These contain *Eco*RI restriction fragments of F cloned into the *Eco*RI site of the plasmid vector pSC101. As the source of R100 DNA, we used pSM6 (4, 6). The single *Bgl*II site of pSM6 is in a region of the plasmid that was expected to lie outside of the regions of homology between F and R100, and so this site was used to cut pSM6 DNA to yield linear molecules. Covalently closed circular pRS26 DNA was digested with *Bam*HI endonuclease, which resulted in two fragments. The smaller fragment contains nearly all of the pSC101 DNA and about half of f12 (7, 10). f12 should not hybridize with pSM6 (8). The larger fragment contains a small piece of pSC101 DNA, all of f15, f1, f17, f19, and f2, as well as the remaining half of f12 (7). Hybridization of this large fragment should yield the desired hybrid molecule with pSM6 cut with *Bgl*II. One such molecule is shown in Fig. 2. The data obtained from measuring such molecules are shown in Table 1 and summarized diagrammatically in Fig. 3A.

The same pSM6 DNA was also hybridized to linear pRS29 DNA (cut at its single *Hin*dIII site in pSC101) and to linear pRS31 DNA (cut at its single *Sal*I site in pSC101). The data obtained using pRS29 and pRS31 are shown in Table 1 and summarized diagrammatically in Fig. 3B and C, respectively.

As expected, the value for hybrids between pSM6 and pRS26 is the sum of the values obtained with pRS29 and pRS31. This enables us to map the *Eco*RI restriction site located at the junction of f1 and f17 (Fig. 1 and 3).

Achtman et al. (1) interpreted genetic data to suggest that one *Eco*RI site in plasmid R6-5 (probably identical to that between fragments r_6 and the r_5-r_8 hybrid of pSM6) maps at a corresponding position to that site between F fragments f1 and f17. The distance between the *Bgl*II site used to linearize pSM6 and the F *Eco*RI site at the junction of f1 and f17 (Fig. 3) is 8.9 ± 0.39 kb (obtained by measuring on 12 molecules of pRS29/pSM6 heteroduplexes the region of nonhomology adjacent to region 5). The distance between the *Bgl*II site and the *Eco*RI site between pSM6 fragments r_6 and the r_5-r_8 hybrid is 9.1 ± 0.35 kb (4). To determine whether the difference in the measurements was real (i.e., the sites are nonhomologous), we analyzed hybrid molecules between pSM6 and pRS29 both digested with *Eco*RI. A typical heteroduplex is shown in Fig. 4. If the *Eco*RI sites were identical, no single-stranded stubs should be detectable extending from the region 5'. None was detected in 10 molecules. Thus, within the error of the measurement of this double-stranded region (±50 base pairs), the *Eco*RI sites in R100-*tra*G and F-*tra*G are identical. This site lies at 82.3 kilobases (kb) on the F map or ~70 kb on the R100 map, taking the R100 coordinate of the

† Present address: Department of Microbiology and Immunology, The University of Adelaide, Adelaide, South Australia 5001.

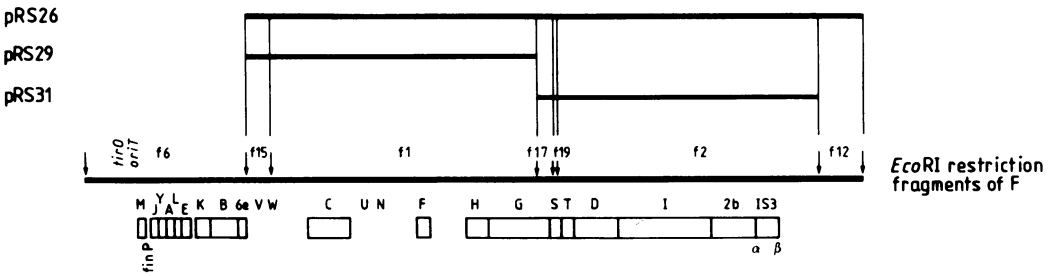

FIG. 1. The F factor *tra* operon and the F DNA encoded by the chimeric plasmids used in this study. The thick lines represent the DNA cloned into pSC101 in the various chimeric plasmids. The boxes represent the amount of DNA required for the various F *tra* proteins. The *traV*, *W*, *U*, and *N* proteins have not as yet been identified.

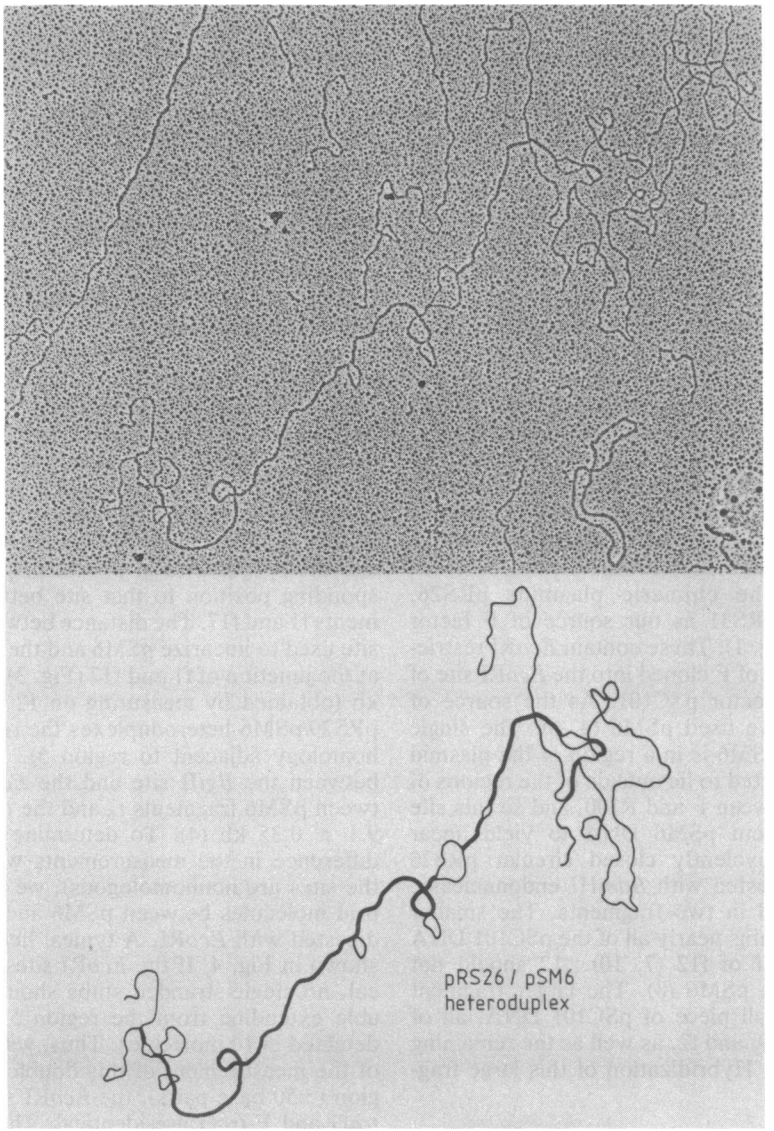

FIG. 2. Heteroduplex of R100 DNA from pSM6 with F DNA from pRS26. The thick lines represent double-stranded DNA, and the thin lines represent single-stranded DNA.

TABLE 1. Measurements of DNA heteroduplexes[a]

Region	pRS26		pRS29		pRS31	
	No. of molecules	Size (kb)	No. of molecules	Size (kb)	No. of molecules	Size (kb)
1	10	0.60 ± 0.03	20	0.70 ± 0.10		
2	10	0.43 ± 0.04	20	0.41 ± 0.04		
3	9	2.03 ± 0.12	16	2.11 ± 0.20		
4	9	0.20 ± 0.04	17	0.15 ± 0.01		
5	8	0.62 ± 0.05				
5'[b]			12	0.20 ± 0.03		
5''[b]					26	0.56 ± 0.07
6	15	0.75 ± 0.07			29	0.72 ± 0.07
7	14	0.80 ± 0.06			27	0.88 ± 0.07
8	16	0.89 ± 0.05			28	0.87 ± 0.11
9	15	2.07 ± 0.11			28	2.08 ± 0.08
10	20	0.86 ± 0.03			27	0.83 ± 0.07
11	20	5.84 ± 0.23			25	5.74 ± 0.24
12	25	1.18 ± 0.08			25	1.14 ± 0.08
13	24	0.40 ± 0.04			25	0.40 ± 0.03
14	22	1.32 ± 0.05			23	1.27 ± 0.08
15	22	0.16 ± 0.02			22	0.16 ± 0.03
16	23	0.34 ± 0.02			22	0.33 ± 0.03
17	23	0.35 ± 0.01			25	0.34 ± 0.02

[a] Sizes were obtained using open circular ColE1 as a standard for double-stranded DNA and φX174 for single-stranded DNA. The final values used for Fig. 3 are the composites of both sets of data presented for each region here.

[b] Regions 5' and 5'' are those parts of region 5 on pRS29 and pRS31, respectively.

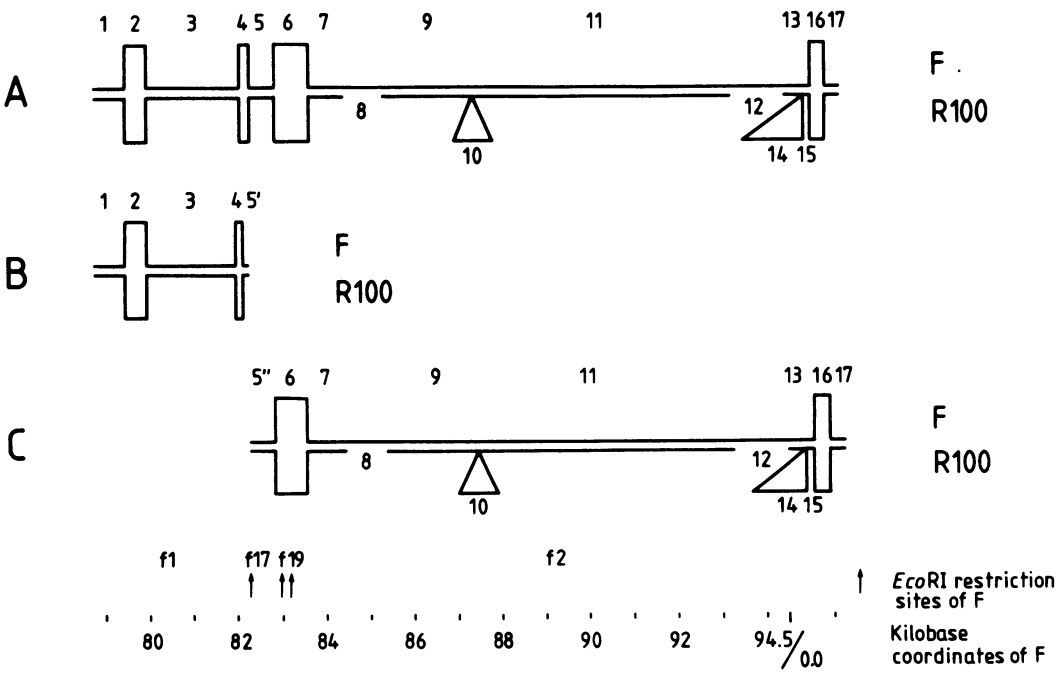

FIG. 3. Diagrammatic representation of the data shown in Table 1. The figures show the heteroduplex structures obtained with R100 DNA, using pSM6 as source, and F DNA, using pRS26 (A), pRS29 (B), or pRS31 (C) as source.

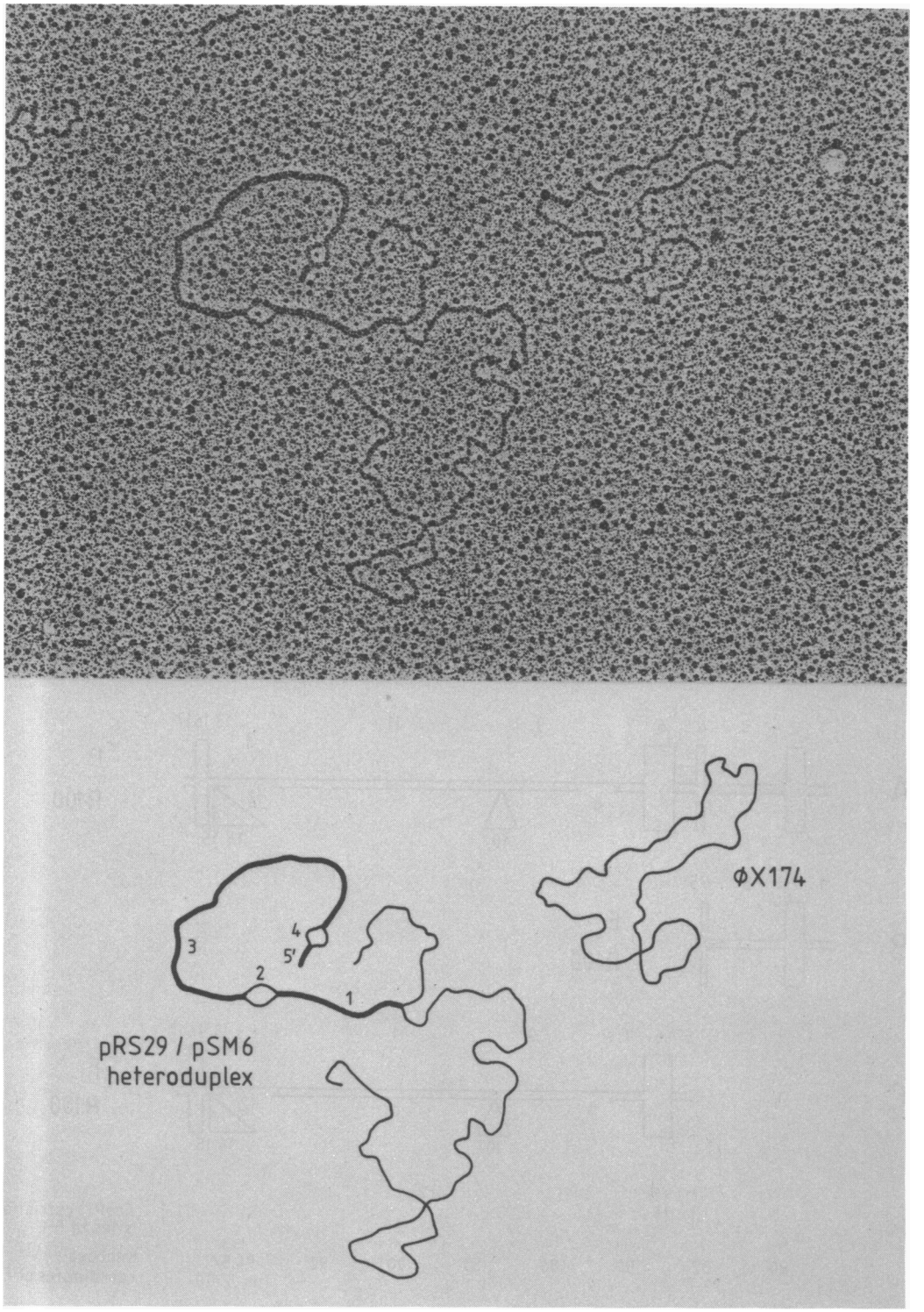

FIG. 4. Heteroduplex structure obtained with pRS29 and pSM6 digested with *Eco*RI.

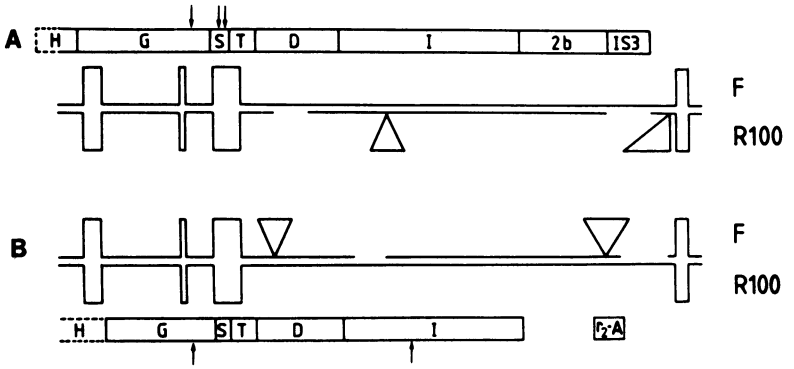

FIG. 5. Heteroduplex analysis of the distal region of the *tra* operon of R100 contained within pSM6 with that of F. (A) Drawn using F DNA as reference. (B) Drawn using pSM6 DNA as reference. Close horizontal parallel lines represent regions of DNA homology (convention of Sharp et al., 8). Distant parallel lines represent substitutions, whereas triangles represent insertions, and spaces represent deletions. IS*3* is absent on pSM6 and is represented as a deletion in pSM6 (A) or insertion in F (B). The arrows represent the *Eco*RI sites in F DNA (A) and in R100 DNA (B). The dashed lines indicating R100-*traH* represent the possible region encoding this gene, although R100-*Tra*Hp has not been identified.

*Eco*RI site between r_2 and the r_5-r_8 hybrid as 87.2 kb.

The data (Fig. 3, Table 1) were also compared with those obtained with intact F′ and R factor DNAs by Sharp et al. (8). All of the insertion, deletion, and substitution loops previously described were detected, and the assignment of each loop as an insertion or deletion by these authors was supported by the new data. Additionally, two small substitution loops (regions 2 and 4) were detected, and more precise data for region 16 were obtained. Where substitutions occurred, the lengths of the unhybridized strands from the two plasmids were indistinguishable.

Using the assignment of region 12 to the IS*3* sequence in F DNA at 93.3 to 94.5 kb (8) and the precise mapping presented here of the *Eco*RI site in *traG* common to both plasmids (see above), the restriction maps of R100 and F could be aligned with the heteroduplex maps and the genetic maps. In this way, it was possible to derive Fig. 5, which we have used as a basis for comparing the physical maps of the two plasmids (4, 5).

The left endpoint of the homologous DNA between pSM6 and either pRS26 or pRS29 is at 78.9 kb (on the F map). This most likely represents one end of the fusion point between R100 *Eco*RI fragments r_5 and r_8, which arose during formation of pSM6. Then 3.4 kb of the fusion fragment would come from r_5, and 2.2 kb would come from r_8. This information has been used to map the *traH* cistron of F and R100 (4, 5).

The hybrid molecules between pRS26 and pSM6 revealed a region of homology between the two which was not contiguous to the DNA regions just described. In several molecules, the left end of pRS26 had hybridized to pSM6 (Fig. 2), as interpreted in Fig. 6 and summarized in Table 2. Assuming pSM6 contains IS*1*, as has been shown for a number of other pSM plasmids (6), we can speculate as to the origins of these regions.

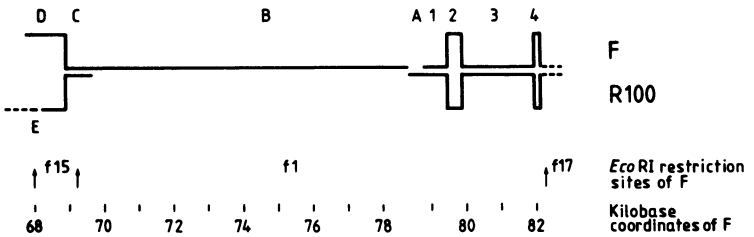

FIG. 6. Diagrammatic representation of the data shown in Table 2, demonstrating the heteroduplex structure obtained with pRS26 and the region within which the fusion between fragments r_5 and r_8 of R100 in pSM6 has occurred. Regions 1, 2, 3, etc., are as for pRS26 in Fig. 3.

TABLE 2. Measurements obtained from the molecules shown diagrammatically in Fig. 6

Region	No. of molecules	Size (kb)
A	10	0.38 ± 0.07
B	—	(9.1)[a]
C	10	0.73 ± 0.05
D	10	1.15 ± 0.07
E	10	5.37 ± 0.51

[a] The size of region B was calculated using values of 1.25 and 13.3 kb for fragments f15 and f1, respectively, and a value of 0.27 kb for the distance of the *Bam*HI site at which pRS26 was cut to the left *Eco*RI site in fragment f15. Electron microscopic values for this large single-stranded DNA region were somewhat variable.

The R100 coordinates of region C are 88.6 to 89.3 kb, which corresponds to the position of the IS*1* sequence from which other pSM plasmids have been generated. There is a low-level promoter in IS*1* capable of expressing the *tra* genes on pSM6 (3). However, the presence of IS*1* in pSM6 would imply that an IS*1* is also present in the F DNA since region C is a homoduplex. The origin of region A is uncertain. The two possibilities are that it represents an insertion in R100 with respect to F DNA or that it is a piece of R100 DNA which became contiguous with the other R100 DNA during the formation of pSM6. This latter possibility implies that pSM6 was formed by an internal deletion from a deletion mutant of R12 (R100), as has occurred with at least one other pSM plasmid (6).

In conclusions, these results show that the similarity of the two sex factors, F and R100, at the level of heteroduplex homology is a useful guide to the extent of their evolution from one another or a common ancestor. This evolution can be taken a step further to the proteins which are encoded within the DNA, and this has been discussed elsewhere (4).

We acknowledge the excellent technical assistance of Claudine Fisseau. We are also most grateful to Mark Achtman for many discussions and for critically reading this paper.

During this study P. A. Manning was supported by grant Ac. 36/2 from the Deutsche Forschungsgemeinschaft.

LITERATURE CITED

1. **Achtman, M., B. Kusecek, and K. N. Timmis.** 1978. *tra* cistrons and proteins encoded by the *Escherichia coli* antibiotic resistance plasmid R6-5. Mol. Gen. Genet. **163**:169–179.
2. **Achtman, M., R. A. Skurray, R. Thompson, R. Helmuth, S. Hall, L. Beutin, and A. J. Clark.** 1978. Assignment of *tra* cistrons to *Eco*RI fragments of F sex factor DNA. J. Bacteriol. **133**:1383–1392.
3. **Dempsey, W. B., and S. McIntire.** 1979. Lambda transducing phages derived from a FinO$^-$ R100::λ cointegrate plasmid: proteins encoded by the R100 replication/incompatability region and the antibiotic resistance determinant. Mol. Gen. Genet. **176**:319–334.
4. **Hansen, B. S., P. A. Manning, and M. Achtman.** 1982. Promoter-distal region of the *tra* operon of the F-like sex factor R100 in *Escherichia coli* K-12. J. Bacteriol. **150**:89–99.
5. **Manning, P. A., B. Kusecek, G. Morelli, C. Fisseau, and M. Achtman.** 1982. Analysis of the promoter-distal region of the *tra* operon of the F sex factor of *Escherichia coli* K-12 encoded by the *Eco*RI restriction fragments f17, f19, and f2. J. Bacteriol. **150**:76–88.
6. **Mickel, S., E. Ohtsubo, and W. Bauer.** 1977. Heteroduplex mapping of small plasmids derived from R-factor R12: *in vivo* recombination occurs on IS1 insertion sequences. Gene. **2**:193–210.
7. **Ohtsubo, H., and E. Ohtsubo.** 1977. Repeated DNA sequences in plasmids, phages, and bacterial chromosomes, p. 49–63. *In* A. J. Bukhari, J. A. Shapiro, and S. L. Adhya (ed.), DNA insertion elements, plasmids, and episomes. Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.
8. **Sharp, P. A., S. N. Cohen, and N. Davidson.** 1973. Electron microscopic heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. II. Structure of drug resistance (R) factors and F factors. J. Mol. Biol. **75**:235–255.
9. **Skurray, R. A., H. Nagaishi, and A. J. Clark.** 1978. Construction and *Bam*HI analysis of chimeric plasmids containing *Eco*RI DNA fragments of the F sex factor. Plasmid **1**:174–186.
10. **Tait, R. C., and H. W. Boyer.** 1978. Restriction endonuclease mapping of pSC101 and pMB9. Mol. Gen. Genet. **164**:285–288.